

Datenaufbereitung aus dem Sozio-ökonomischen Panel: Erwerbsverlauf und Familiendaten

Ernicke, Beate

Veröffentlichungsversion / Published Version
Arbeitspapier / working paper

Empfohlene Zitierung / Suggested Citation:

Ernicke, B. (1997). *Datenaufbereitung aus dem Sozio-ökonomischen Panel: Erwerbsverlauf und Familiendaten*. (Arbeitspapier / Sfb 186, 44). Bremen: Universität Bremen, SFB 186 Statuspassagen und Risikolagen im Lebensverlauf. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-57503>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC-ND Lizenz (Namensnennung-Nicht-kommerziell-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de>

Terms of use:

This document is made available under a CC BY-NC-ND Licence (Attribution-Non Commercial-NoDerivatives). For more information see:
<https://creativecommons.org/licenses/by-nc-nd/4.0>

**Sonderforschungsbereich 186
der Universität Bremen**

**Statuspassagen und Risikolagen
im Lebensverlauf**

**Datenaufbereitung aus dem
Sozio-ökonomischen Panel:
Erwerbsverlauf und Familiendaten**

von

Beate Ernicke

Arbeitspapier Nr. 44
Bremen, Februar 1997



**Sonderforschungsbereich 186
der Universität Bremen**

**Statuspassagen und Risikolagen
im Lebenslauf**

**Datenaufbereitung aus dem
Sozio-ökonomischen Panel:
Erwerbsverlauf und Familiendaten**

**Beate Ernicke
Wiener Straße
Postfach 330440
28334 Bremen**

**Tel.: 0421/218 3298
Fax: 0421/218 7474
e-mail: bernicke@sfb186.uni-bremen.de**

Vorwort

Innerhalb des Sonderforschungsbereiches "Statuspassagen und Risikolagen im Lebenslauf" untersucht das Projekt B6 die Beziehungen der individuellen Lebensläufe zu den strukturellen Veränderungen auf der Makroebene. Zur Realisierung von Forschungsarbeiten in diesem, wie auch in anderen Projekten, besteht ein kontinuierlicher Bedarf an adäquaten Datensätzen. Im vorliegenden Arbeitspapier wird auf die Entstehung einiger der erforderlichen Datensätze eingegangen.

Nach einer Einleitung wird im zweiten Abschnitt das sozio-ökonomische Panel als Grundlage der Datenaufbereitung beschrieben. Der dritte und vierte Abschnitt konzentriert sich auf die Generierung der Datensätze und den dabei auftretenden Fragen.

Für die Lebenslaufforschung sind Biographiedaten über den Erwerbsstatus und den Familienstand (Partner, Ehe, Kinder) aus dem sozioökonomischen Panel von Bedeutung. Sie liefern für Fragestellungen im Sfb wichtige Ereignisdaten, die den Wandel von Biographiemustern dokumentieren.

Prof. Dr. Walter R. Heinz
Sprecher des SFB 186

Inhaltsverzeichnis

1	Einleitung	3
2	Das sozio-ökonomische Panel	5
2.1	Die Erhebung der Biographiedaten	7
2.2	Fehlende Daten	8
3	Daten zu den Erwerbsstatus der Personen	10
3.1	Anforderungen an den Datensatz <i>erwerb_spells.dat</i>	10
3.2	Die Ausgangs-Datensätze im SOEP	12
3.2.1	<i>ap.dat</i> bis <i>kp.dat</i>	12
3.2.2	<i>artkalen.dat</i>	12
3.2.3	<i>pbiospe.dat</i>	14
3.3	Aufbereitung der Datensätze <i>artkalen.dat</i> und <i>pbiospe.dat</i>	16
3.3.1	Anpassung von <i>artkalen.dat</i>	17
3.3.2	Anpassung von <i>pbiospe.dat</i>	18
3.4	Zwischenstand	19
3.5	Überschneidungen und Lücken	20
3.6	Einige Informationen zu <i>erwerb_spells.dat</i>	24
4	Daten zu Ehen und Kindern	28
4.1	Anforderungen an den Datensatz <i>ereignisse.dat</i>	28
4.2	Partnerschaften, Partner, Partnerzeiger	31
4.2.1	Ehedaten aus <i>biolela.dat</i>	32
4.2.2	Ehedaten aus dem 11-Jahres-Fenster	35
4.2.3	Zusammenstellung der Ehe-Ereignisse	37
4.3	Kinder	40
4.3.1	Retrospektiv erhobene Daten über Kinder	41
4.3.2	Geburten im 11-Jahres-Fenster	42
4.4	Zusammenführung der Ereignisdaten	43

A	Statistische Kennzahlen zur Datei <i>erwerb.dat</i>	45
A.1	<i>artkalen.cf</i>	46
A.2	<i>pbiospe.cf</i>	47
A.3	<i>artkalen_spells.dat</i>	48
A.4	<i>pbiospe_spells.dat</i>	52
B	Statistische Kennzahlen zur Datei <i>erwerb_spells.dat</i>	56
C	Statistische Kennzahlen zur Datei <i>ehen.dat</i>	64
C.1	<i>ehen.cf</i>	65
C.2	<i>ehen.dat</i>	67
D	Statistische Kennzahlen zur Datei <i>kinder.dat</i>	84
D.1	<i>kinder.cf</i>	85
D.2	<i>kinder.dat</i>	86
E	Statistische Kennzahlen zur Datei <i>ereignisse.dat</i>	97
F	Literatur	104

Kapitel 1

Einleitung

Die Datenaufbereitung stellt bekanntlich einen Großteil der anfallenden Arbeit bei der empirischen Längsschnittforschung dar, weshalb stets nach bereits aufbereiteten Daten oder nach 'Tips und Tricks' gesucht werden sollte, um die Arbeit der Datenaufbereitung zu minimieren. Stehen keine bereits aufgearbeiteten Daten zur Verfügung, steht der Forscher bei einer neu zu beginnenden Analyse vor der Aufgabe aus der Datenfülle diejenigen Daten herauszufiltern und aufzubereiten, die für seine Analysen benötigt werden. Jeder, der die Aufgabe einer Datenaufbereitung mit Daten aus dem Sozio-ökonomischen Panel (SOEP) einmal übernommen hat, kennt viele der Problematiken bzw. Tücken, die mit dieser Aufgabe verbunden sind. Die anfallenden Arbeiten gestalten sich einfacher, wenn bereits Erfahrungen oder Informationen vorliegen. Dieses soll das vorliegende Arbeitspapier leisten, indem der Weg der Datenaufbereitung ausgehend von einigen Datensätzen des SOEP bis hin zu den fertigen Datensätzen des Projektes B6 aufgezeigt wird. Dabei sollen die aufgetretenen Probleme zusammen mit den Überlegungen zu deren Lösung und schließlich den Entscheidungen dargelegt werden.

Das SOEP besteht aus einer Reihe von Datensätzen, die jeweils Informationen zu einem bestimmten Bereich und/oder Jahr beinhalten. Je nachdem, welche Forschungsfrage bearbeitet werden soll, müssen Informationen aus einem oder mehreren Datensätzen in die gewünschten Ergebnisdatensätze überführt werden. Zur Unterstützung dieser Arbeit gibt es Informationen über die Ausgangsdatsätze des SOEP, die sich zumeist auf eine Beschreibung der einzelnen Spalten und deren Werte beschränken. Ausgestattet mit diesen Informationen beginnt dann die Datenaufbereitung.

Abgesehen davon, daß für eine Forschungsfrage nie die 'perfekten' Daten vorhanden sind, stößt der Forscher sehr schnell bei der Datenaufbereitung auf Probleme. Offensichtlich ist bereits das Problem der kontinuierlichen Teilnahme am SOEP. Es gibt viele Personen, die an einer oder mehr Erhebungen nicht teilgenommen haben und daher zu diesen Jahren keine Daten liefern. Was soll

mit diesen Personen geschehen? Ein Ausschluß dieser Personen löst das Problem radikal, stellt den Forscher aber vor die Frage: Habe ich immer noch eine angemessene Stichprobe, oder sind diese Personen systematische, d. h. "nicht zufällige" Ausfälle?

Wenn man das Ziel hat, eine maximale Anzahl an Personen mit brauchbaren Informationen in die Datensätze aufzunehmen, dann sind eine Reihe von Fragen zu lösen. Manchmal waren die Probleme recht einfach zu lösen, manchmal mußten radikalere Entscheidungen getroffen werden. Im Laufe der Datenaufbereitung wurden viele Schritte mehrmals wiederholt, wenn sich z.B. zeigte, daß getroffene Entscheidungen doch nicht das Optimum darstellten, oder ein Zusammenhang mit anderen Problemen entdeckt wurde. Bei aller Liebe zum Detail konnten aber nicht alle kleinen Probleme, die im Zuge der Arbeit mit dem SOEP auftauchten, im vorliegenden Arbeitspapier ausführlich behandelt werden.

Als Ergebnis der Arbeiten werden zwei Datensätze zur Verfügung gestellt:

1. Ein Datensatz¹ zu den Erwerbsstatus der Personen (*erwerb_spells.dat*): Innerhalb dieses Datensatzes erfolgte die Rekonstruktion der Erwerbskarriere der Personen.
2. Ein Datensatz mit den Daten zu Ehen und Kindern (*ereignisse.dat*).

Der Umfang und die Auswahl der Daten ist bestimmt durch das Projekt "Haushaltsdynamik und soziale Ungleichheit im internationalen Vergleich", welches an der Bremer Universität im Sonderforschungsbereich 186 (Sfb186) als Projekt B6 unter der Leitung von Prof. Dr. Hans-Peter Blossfeld durchgeführt wird.

Die generierten Datensätze sind von grundlegender Bedeutung für die Forschungsarbeit im Projekt B6. Sie bilden bei vielen Fragestellungen die Datengrundlage. Forschungsarbeiten, die diese Datensätze verwendet haben, sind z.B.:

- Blossfeld, Hans-Peter; Timm, Andreas 1996: Soziale Schließung oder Öffnung: Homogamieprozesse in Deutschland. Bremen. (mimeo).
- Drobnič, Sonja; Blossfeld, Hans-Peter; Rohwer, Götz; Wittig, Immo 1996: Women's Part-time Work: Reconciliation Between Family and Employment?. Bremen. Sfb Arbeitspapier Nr 32.

Datenaufbereitung und statistische Analysen erfolgten überwiegend mit dem von Götz Rohwer entwickelten Statistikprogramm TDA (Transitions Data Analysis, Rohwer 1994a). Einige darüber hinaus erforderliche C-Programme wurden auf Unix-Workstations entwickelt.

¹Ein Datensatz entspricht einer Datei.

Kapitel 2

Das sozio-ökonomische Panel

Die erste Erhebung des SOEP war im Jahre 1984 und seitdem erfolgte jährlich eine Erweiterung um eine Welle. Die Erhebungen beschränkten sich zuerst auf die damalige Bundesrepublik Deutschland und wurden ab 1990 auf die damalige Deutsche Demokratische Republik erweitert. Eine weitere Ausdehnung erfolgte 1994 auf Zuwanderer nach Deutschland.

Das Panel ist ausgerichtet auf die Erhebung von Haushalten. Dies bedeutet, daß alle in einem Haushalt lebenden Personen erfaßt werden (Haushaltsebene). Innerhalb eines Haushaltes wird mit jeder Person, die das 16. Lebensjahr erreicht hat, zusätzlich eine personenbezogene Befragung durchgeführt (Individualebene). Für jede der beiden Ebenen wurde ein Fragebogen entwickelt: der Haushaltsfragebogen, der von einer Person des Haushaltes, vorzugsweise dem Haushaltsvorstand, beantwortet werden sollte und der Personenfragebogen für jede im Haushalt lebende Person, die das 16. Lebensjahr erreicht hat.

Die Zuordnung einer Person zum Haushalt und den darin lebenden anderen Personen bleibt erhalten und ist damit rekonstruierbar. Dieses ist wichtig, wenn z.B. Partnerschaften erforscht werden sollen.

Die Stichprobe der Haushalte umfaßte anfänglich zwei Teile:

- Personen in westdeutschen Privathaushalten, deren Haushaltsvorstand nicht die türkische, griechische, jugoslawische, spanische oder italienische Staatsangehörigkeit besitzt (Stichprobe A)¹
- Personen in westdeutschen Privathaushalten, deren Haushaltsvorstand die türkische, griechische, jugoslawische, spanische oder italienische Staatsangehörigkeit besitzt (Stichprobe B).

Ergänzt wurde diese Liste 1990 um die Stichprobe C, Personen von Haushalten aus den neuen Bundesländern. Hier fand keine Unterteilung nach deutschem oder

¹Die Staatsangehörigkeit kann über ein Datum, welches die Nationalität repräsentiert, rekonstruiert werden.

ausländischem Haushaltsvorstand statt. 1994 erfolgte eine Erweiterung um die Stichprobe der Zuwanderer nach Deutschland (Stichproben D1 und D2).

Innerhalb des SOEP wurden und werden eine Reihe von retrospektiven Daten erhoben:

- das Kalendarium:

Mit dem Fragenkomplex des Kalendariums werden die Daten für das Vorjahr und das gegenwärtige Jahr bis zur Erhebung bezüglich Erwerbsstatus und Einkommensquellen erfaßt. Diese Daten werden in jeder Welle erhoben.

- das Biographieschema:

Mit dem Fragenkomplex des Biographieschemas werden die Daten zu Erwerbsstatus und Einkommensquellen vom 15. Lebensjahr an bis zu dieser Erhebung, altersbezogen und auf Jahresbasis, retrospektiv erfaßt. Diese Befragung findet für jede Person nur einmal statt. Der Zeitpunkt dieser Erhebung war für die Teilnehmer der ersten Welle die erste, aber für hinzukommende Personen spätere Erhebungswellen, wobei die Erhebung nicht mit dem Zeitpunkt der Aufnahme in das SOEP übereinstimmt².

- weitere Biographiedaten:

Zuzüglich zum Biographieschema werden weitere Daten gleichzeitig ereignisorientiert erhoben. Dies sind demographische und erwerbsbiographische Daten. Auch diese Befragung findet nur einmal für jede Person statt. Im vorliegenden Arbeitspapier fassen wir diese Daten und das oben genannte Biographieschema unter dem Oberbegriff Biographiedaten zusammen.

- weitere jährliche Erhebungsdaten:

Jährlich erhoben werden darüber hinaus einige demographische und erwerbsbiographische Daten, die sich auf Veränderungen der Lebenssituation seit der letzten Befragung beziehen.

Im vorliegenden Arbeitspapier werden die Begriffe Kalendarium, Biographieschema und Biographiedaten entsprechend diesen Erläuterungen verwendet. Für die Arbeit lagen die Daten der ersten 11 Erhebungswellen vor. Diese 11 Wellen bezeichne ich im Folgenden als 11-Jahres-Fenster.

Für die weitere Arbeit sind die Biographiedaten von besonderer Bedeutung, da sie den Lebenslauf einer Person bis zur Biographie-Erhebung umfassen. Im folgenden Abschnitt wird kurz auf die Erhebung der Biographiedaten eingegangen.

²Dadurch sind Überschneidungen der retrospektiv erhobenen Daten mit den Daten aus den Panelwellen möglich.

2.1 Die Erhebung der Biographiedaten

Die Biographiedaten im SOEP umfassen folgende Datenbereiche:

- das Biographieschema, d.h. der Erwerbsstatus,
- die Kinder- und Ehebiographie,
- die soziale Herkunft und
- die Daten zum Berufseinstieg.

Die Daten des Biographieschemas liegen jahresbezogen vor. Für jedes Jahr ist festgehalten, welchen Erwerbszustand bzw. -status eine Person angenommen hat. Die Daten zu den Erwerbsstatus der Personen wurden einmal retrospektiv erfaßt. Dabei waren die Befragten aufgefordert, Angaben über ihre bisherige Erwerbskarriere zu machen. Dies bestand aus Ankreuzen von Kästchen, die "Zuständen" vom 15. bis zum 65. Lebensjahr, bzw. für Personen, die jünger als 65 Jahre sind, bis zum Erhebungsjahr entsprachen. Folgende Zustände wurden anfänglich³ unterschieden:

- in Schule, Studium, Abendschule
- in Lehre, Berufsausbildung, Fortbildung, Umschulung
- beim Wehrdienst / Zivildienst, im Krieg, in Gefangenschaft
- voll berufstätig (einschl. Zeit- / Berufssoldat)
- teilzeitbeschäftigt oder geringfügig erwerbstätig
- arbeitslos gemeldet
- Hausfrau, Hausmann
- im Ruhestand, Rentner / Pensionär
- Sonstiges

Diese retrospektiven Daten wurden erstmalig 1984 für alle Teilnehmer am SOEP erhoben; der Fragenkomplex befand sich innerhalb des Personenfragebogens. In den nächsten zwei Wellen 1985 und 1986 wurde dieser Fragenkomplex nicht erhoben, statt dessen erfolgte 1985 eine Biographieerhebung zum Thema

³Dieser "Zustandsraum" wurde für westdeutsche Befragte 1991 (8.Welle) um den Zustand "Im Mutterschafts- / Erziehungsurlaub" und 1992 (9.Welle) um den Zustand "In Kurzarbeit oder Warteschleife" erweitert. Die zweite Erweiterung ergab sich u.a. aus der Zusammenführung der Befragungen West- und Ostdeutschlands in einen gemeinsamen Fragebogen.

"Kinder- und Ehebiographie" und 1986 zum Thema "Berufseinstieg und soziale Herkunft". In der vierten Welle 1987 fand keine Biographieerhebung statt.

1988 wurde aus den drei Biographiekomplexen Biographieschema (von 1984), Kinder- und Ehebiographie (von 1985) sowie soziale Herkunft und Berufseinstieg (von 1986) ein gemeinsamer Biographiekomplex zusammengestellt, der als Ergänzung des Personenfragebogen für alle erstmals teilnehmenden Personen herangezogen worden war. Diese Zusatzbefragung wurde in den Wellen 5 bis 7 (1988 bis 1990) durchgeführt. Ab 1991 erfolgt die Erhebung der Biographiedaten in einem separaten Biographiefragebogen als Zusatzbefragung. 1992 fand eine erste und 1993 eine zweite Nacherhebungen für Personen statt, bei denen die Biographiedaten fehlten. Seit 1992 wird die Biographieerhebung als Zusatzbefragung jährlich für alle neu hinzukommenden Personen herangezogen. Seit 1994 steht dazu ein modifizierter Fragebogen "Lebenslauf" zur Verfügung. Einen Überblick über die Biographieerhebung der 11 Wellen des SOEP liefert die Tabelle 2.1 (vgl. Landua 1995).

Tabelle 2.1: Überblick über die Biographiedatenerhebungen
Biographiedatenerhebungen

1. Welle 1984	Erhebung des Biographieschemas für die Teilnehmer an der ersten Welle
2. Welle 1985	Erhebung der Kindheits- und Ehebiographie für die Teilnehmer an der zweiten Welle
3. Welle 1986	Erhebung der sozialen Herkunft und des Berufseinstiegs für die Teilnehmer an der dritten Welle
4. Welle 1987	
5. Welle 1988	Erhebung des gesamten Biographiebereiches für die neuen Teilnehmer
6. Welle 1989	Erhebung des gesamten Biographiebereiches für die neuen Teilnehmer
7. Welle 1990	Erhebung des gesamten Biographiebereiches für die neuen Teilnehmer (1. Welle für Ostdeutschland)
8. Welle 1991	Erhebung des gesamten Biographiebereiches für die neuen Teilnehmer, gesonderter Fragebogen
9. Welle 1992	Erhebung des gesamten Biographiebereiches für die neuen Teilnehmer und Nacherhebung in Westdeutschland
10. Welle 1993	Erhebung des gesamten Biographiebereiches für die neuen Teilnehmer und Nacherhebung in Ostdeutschland
11. Welle 1994	Erhebung des gesamten Biographiebereiches für die neuen Teilnehmer und Nacherhebung in Ostdeutschland, modifizierter Fragebogen "Lebenslauf"

2.2 Fehlende Daten

Es ist leicht ersichtlich, daß von einer Person nur dann die jährlichen Daten vorliegen, wenn die Person zur Zeit der Erhebung zur Verfügung stand. Andernfalls

besitzt die Person für dieses Jahr bzw. für das vorherige Jahr eine Lücke in den Daten. Darüber hinaus gibt es das Problem eines "fehlenden Lebenslaufes", wenn von einer Person keine Daten der Biographieerhebung vorliegen.

Im SOEP hat der Forscher es mit beiden Problematiken zu tun. Die Nichtteilnahme von Personen an einer oder mehreren Wellen führt zu teilweise erheblichen Lücken in den Daten. Die Problematik mit den fehlenden Biographiedaten sollte durch die Nacherhebungen ab 1992 beendet sein, doch standen diese Daten bislang nicht zur Verfügung.

Mit den jetzt vorliegenden zwei neuen und wichtigen Datensätzen *pbiospe.dat* und *biolela.dat* können die Daten aus den Nacherhebungen für die Forschung genutzt werden. In *pbiospe.dat* stehen die Daten zum Biographieschema aufbereitet als Ereignisdaten zur Verfügung. Alle weiteren Biographiedaten befinden sich in *biolela.dat*.

Kapitel 3

Daten zu den Erwerbsstatus der Personen

Lebenswege von Personen lassen sich unter den unterschiedlichsten Gesichtspunkten gliedern. Das Projekt B6 benötigt zumeist eine Beschreibung von Lebenswegen anhand des Erwerbsverhaltens. Liegt dieser Lebensweg in Spelldatenformat vor, so bildet der entstandene Datensatz die Grundlage für Längsschnittanalysen.

Dieses Kapitel beschreibt, wie ausgehend von einigen Datensätzen des SOEP ein solcher Datensatz gewonnen wurde. Dazu erläutere ich als erstes, welche Anforderungen an den Ergebnisdatensatz gestellt wurden, damit dieser den Bedürfnissen des Projektes B6 gerecht werden konnte. Anschließend gehe ich auf die unterschiedlichen Ausgangsdatsätze des SOEP ein, betrachte diese genauer und wähle erforderliche Spalten aus. Durch schrittweises anpassen der Daten an die Anforderungen des Projektes B6 entsteht schließlich die Ergebnisdatei *erwerb_spells.dat*.

3.1 Anforderungen an den Datensatz *erwerb_spells.dat*

Der Datensatz mit den Erwerbsstatus der Personen stellt für Längsschnittanalysen eine wichtige Grundlage dar. Viele Fragestellungen bauen auf dieser Art der Darstellung von Lebensläufen auf. Als Erwerbsstatus werden in dieser Datei alle Zustände aufgenommen, die eine Lebenssituation einer Person beschreiben, wie "in der Schule", "Hausfrau und -mann" oder "Rentner". Es sind also keineswegs nur Zustände gemeint, die mit einer Art von Einkommen verbunden sind. Möglich wäre als Erweiterung der Zustände die Aufnahme von Mehrfach-Zuständen, da sich im Laufe der Arbeit mit dem SOEP herausgestellt hat, daß

z.B. einige Personen gleichzeitig vollzeitbeschäftigt sind und eine Schule besuchen. Neben sinnvollen Kombinationen, wie die oben genannte, sind eine Reihe von nicht sinnvollen Kombinationen zu erkennen, z.B. die Verbindung von Vollzeitbeschäftigung und Arbeitslosigkeit.

Das Projekt B6 hat sich für Einzelzustände entschieden und eine Rangfolge festgelegt, welcher Zustand bei einer Mehrfachnennung übernommen werden soll:

1. Vollzeitbeschäftigung
2. Teilzeitbeschäftigung
3. berufliche Ausbildung
4. Arbeitslosigkeit
5. im Ruhestand
6. in Schule, Hochschule, Universität
7. bei Wehr- oder Zivildienst
8. Hausfrau oder -mann
9. sonstiges

Eine Verringerung des Zustandsraumes ist in jeder späteren Phase eines Projektes möglich, nicht jedoch die Erweiterung, da die notwendigen Informationen dann nicht mehr zur Verfügung stehen. Die Entscheidung welcher Zustandsraum angemessen ist, steht in enger Verbindung zur Forschungsfrage. Als Beispiel sei hier erwähnt, da Personen, die sich im Mutterschaftsurlaub befinden und von ihrer Arbeit freigestellt sind, scheinbar (ebenfalls) Vollzeit- bzw. Teilzeitbeschäftigung ankreuzen und somit in diese Gruppen aufgenommen werden.

Der angestrebte Datensatz *erwerb_spells.dat* soll in seiner Struktur einen Spelldatensatz darstellen. Ein Spell entspricht dabei einer Erwerbsperiode, d.h. der Darstellung eines Zeitraumes, in dem eine Person einen Zustand angenommen hat. Ein Beispiel eines Spells ist: die Person 3645 habe im Zeitraum vom 1.1.94 bis zum 31.4.94 eine Teilzeitbeschäftigung ausgeübt. Im späteren Datensatz erscheinen die Daten allerdings nicht in dieser sprachlichen Form, sondern kodiert als Zahlen.

Für diesen Datensatz sind nachstehend genannte Informationen zu verbinden:

- Jede Person wird eindeutig durch die Personenummer identifiziert. Weiterhin gehen für die Personen das Geschlecht, das Geburtsjahr und die Stichprobenzugehörigkeit als Grundvariablen in den Datensatz ein.
- Jeder Spell entspricht einer Zeile bzw. einem Rekord in der Datei *erwerb_spells.dat*.

- Jeder Spells erhält eine Nummer, die aufsteigend von 1 an für jede Person die Spells nummeriert.
- Für jeden Spell sind der Erwerbszustand der Person innerhalb des Spells, der Folgezustand, der Anfangs- und der Endzeitpunkt des Spells aufzunehmen.
- Für jeden Spell ist die Art der Zensierung¹ aufzuzeigen.
- Für jeden Spell ist die Dauer des Spells zu berechnen.

Mit diesen Daten ist die Grundstruktur des Datensatzes *erwerb_spells.dat* dargelegt.

3.2 Die Ausgangs-Datensätze im SOEP

Die zu generierenden Erwerbsspellss werden die Grundlage späterer Analysen bilden und Informationen zu den Erwerbsphasen der Personen bereithalten. Daten zur Erwerbstätigkeit der Personen liegen in unterschiedlicher Form im SOEP vor: Neben den Rohdaten in *ap.dat*, *bp.dat* bis *kp.dat* liegen zwei Dateien mit bereits aufgearbeiteten Daten vor, die Dateien *artkalen.dat* und *pbiospe.dat*. Beide Dateien liefern eine Art Spelldaten, allerdings in einer zueinander unterschiedlichen Form. Die Dateien *ap.dat* bis *kp.dat*, *artkalen.dat* und *pbiospe.dat* sind im normalen Lieferumfang des SOEP enthalten.

3.2.1 *ap.dat* bis *kp.dat*

Die Dateien *ap.dat*, *bp.dat* bis *kp.dat* entsprechen den Daten aus den Befragungen der einzelnen Wellen. Die Datei *ap.dat* beinhaltet die Daten der ersten Welle, *bp.dat* die Daten der zweiten Welle usw., bis zur Datei *kp.dat* mit den Daten der elften Welle. Diese Daten können zur Erstellung der angestrebten Dateien herangezogen werden, doch liegen dieselben Daten bezgl. der Erwerbskarriere in der Datei *artkalen.dat* bereits etwas aufbereitet vor.

3.2.2 *artkalen.dat*

In *artkalen.dat* sind die Erwerbsspellss des 11-Jahres-Fensters² als Spelldaten aufbereitet. In der Tabelle 3.1 wird der Aufbau der Datei dargestellt.

¹Zensierungen sind Rechts-, Linkszensierung und die beidseitige Zensierung.

²Es ist zu beachten, daß die Erhebung eines Jahres u.a. Daten des vorherigen Jahres liefert.

Tabelle 3.1: Aufbau der Datei *artkalen.dat*

Spalte 1:	Nummer des Ursprungshaushaltes
Spalte 2:	Personennummer
Spalte 3:	Laufende Nummer des Ereignisses bezogen auf die Person
Spalte 4:	Art des Ereignisses (Zustand einer Person); diese bestimmen in ihren Ausprägungen den Zustandsraum.
Spalte 5:	Anfangsmonat des Ereignisses
Spalte 6:	Endmonat des Ereignisses; Anfangs- und Endmonat sind als Zahlen von 1 bis 132 gegeben, wobei 1 den Januar 1983 und 132 den Dezember 1993 bezeichnet.
Spalte 7:	Zensierungsspalte; Die Art der Zensierung jedes Spells.
Spalte 8:	Stichprobenzugehörigkeit; Über die Stichprobenzugehörigkeit erfolgt später die Eingrenzung auf Personen der Stichproben A und B (vgl. Seite 4)

Sie beinhaltet insgesamt 67170 Rekords, wobei ein Rekord eine Zeile darstellt. Interessant ist die Spalte 4: Spelltyp oder Zustand. Ihre Ausprägungen sind, zusammen mit einer Häufigkeitsauszählung, in der Tabelle 3.2 aufgelistet.

Tabelle 3.2: Ausprägungen der Zustände in *artkalen.dat*

Zustand	Anzahl	Kurzbeschreibung
1	20874	Voll erwerbstätig
2	1593	Kurzarbeit
3	6419	Teilzeit-geringfügig
4	4699	Betriebliche Ausbildung
5	6740	Arbeitslos gemeldet
6	4909	in Rente - Ruhestand
7	1533	Mutterschafts-Freistellung
8	5552	Schule - Hochschule
9	753	Wehr-, Zivildienst
10	8482	Hausfrau, -mann
11	514	Nebenberufliche Tätigkeit
12	2845	sonstiges
99	2257	unbekannt

Dort ist erkennbar, da es 2257 Spells mit dem Zustand 99 gibt. Diese Spells besitzen zugleich in der Spalte Zensierung den Wert -2. Der Zustand 99 ist ein unbekannter Zustand³, der diese Spells für das Anliegen unbrauchbar macht. Die Spells werden daher aus dem Datensatz entfernt.

Die Spalte 5 (Anfangsmonat/Spellanfang) zeigt einige Besonderheiten:

- Innerhalb eines Jahres fangen viele Spells im Januar an, was durch die auf Jahresbasis⁴ erhobenen Daten verursacht wird.

³Mir war es nicht möglich diesem Zustand eine Bedeutung zuzuordnen.

⁴Die Daten wurden, wie bereits oben geschrieben alterbezogen vom 15. bis zum 65. Lebensjahr auf Jahresbasis erhoben.

- Die Spellstarts zu 79 (Juli 1989) und 91 (Juli 1990) sind überdurchschnittlich stark besetzt mit 5105 Spells bzw. 1018 Spells. Dieses steht in Verbindung mit der Erhebung in Ostdeutschland und kann daher, da ostdeutsche Personen nicht einbezogen werden sollen, unbeachtet bleiben. Bei der späteren Selektion reduziert sich die Anzahl der Spells auf ein "normales" Maß⁵.

Bei der Spalte 6 (Endmonat/Spellende) zeigt sich ein ähnliches Bild:

- Innerhalb eines Jahres enden viele Spells im Dezember, was wiederum durch die auf Jahresbasis erhobenen Daten verursacht wird.
- Der Monat 90 (Juni 1990) und der Monat 99 (März 1991) sind neben dem Dezember überdurchschnittlich stark besetzt mit 1397 bzw. 1201 Spells. Dies steht wie bei dem oben genannten Spellanfang in Verbindung mit der Erhebung in Ostdeutschland und kann daher unbeachtet bleiben.

Von den Spalten aus *artkalen.dat* verwende ich nur die Spalten 2, 4, 5 und 6 (Personennummer, Art des Ereignisses, Anfangsmonat und Endmonat). Die Stichprobenzugehörigkeit, die für Selektionen notwendig ist, wird später aus einer anderen Datei hinzugenommen⁶.

3.2.3 *pbiospe.dat*

Die Datei *pbiospe.dat* vereinigt in sich die Erwerbsdaten der Biographieerhebungen und des 11-Jahres-Fensters. Da die Start- und Endzeitpunkte der Spelldaten auf Jahresbasis festgelegt wurden, gehen die monatsbezogenen Informationen der in den einzelnen Wellen erhobenen Daten verloren. Dieser Informationsverlust kann etwas ausgeglichen werden, indem aus dieser Datei nur die Spells bis zu dem Jahr herangezogen werden, bis die monatsbezogenen Daten aus den einzelnen Wellen vorliegen. Von da an greife ich auf die Daten aus *artkalen.dat* zurück; hier liegen die Anfangs- und Endzeitpunkte monatsbezogen vor. Der Aufbau der Datei *pbiospe.dat* ist in der Tabelle 3.3 dargestellt.

⁵Zur Verteilung der Spells zu den Stichproben:

Stichprobe	Spellanfang 79	Spellanfang 91	Spellende 90	Spellende 99
A	169	169	179	224
B	55	55	60	63
C	4881	794	1158	914

⁶Ich verwende nicht die Stichprobenzugehörigkeit aus *artkalen.dat*, da ich ohnehin weitere Daten aus der Datei (*ppfad.dat*) benötige und es sich gezeigt hat, daß diese Datei in Bezug zur Güte der Daten anderen Dateien oft überlegen, aber zumindest gleich ist.

Tabelle 3.3: Aufbau der Datei *pbiospe.dat*

Spalte 1:	Ursprungs Haushaltsnummer
Spalte 2:	eindeutige Personennummer
Spalte 3:	fortlaufende Spellnummer, bezogen auf eine Person
Spalte 4:	Spelltyp; Zustand einer Person
Spalte 5:	Beginn des Spells
Spalte 6:	Ende des Spells
	Beginn und Ende liegen jahresbezogen vor, wobei die Jahreszahlen dem Alter der Person entsprechen.
Spalte 7:	Startjahr des Anteils des Spells, der aus dem Biographieschema stammt
Spalte 8:	Endjahr des Anteils des Spells, der aus dem Biographieschema stammt
Spalte 9:	Startjahr des Anteils des Spells, der aus dem Kalendarium stammt
Spalte 10:	Endjahr des Anteils des Spells, der aus dem Kalendarium stammt
Spalte 11:	Zensierung für kombinierte Spells, mit: 1 = nicht zensiert, 2 = links zensiert, 3 = rechts zensiert, 4 = links und rechts zensiert.
Spalte 12:	Information zum Spell In dieser Spalte wird aufgezeigt, aus welchen Teilen sich der Spell zusammensetzt. Dieser kann rein aus der Biographieerhebung stammen, rein aus dem Kalender stammen, ein kombinierter Spell aus Biographie und Kalendarium sein und geschlossene Lücken beinhalten.
Spalte 13:	Erhebungsjahr des Biographieschemas
Spalte 14:	Fehlercode
Spalte 15:	Stichprobenzugehörigkeit 1 = Stichprobe A ; 2 = Stichprobe B ; 3 = Stichprobe C;

Für die Aufbereitung der Daten sind, wie bereits bei *artkalen.dat* angedeutet, nicht alle Spalten von Bedeutung, weshalb innerhalb dieses Arbeitspapiers auch nur eingehender auf die Spalten eingegangen wird, die für die weitere Arbeit notwendig sind. Dies sind innerhalb der Datei *pbiospe.dat* die Spalte 2, 4, 5 und 6.

An dieser Stelle möchte ich auf zwei Fehlercodes hinweisen, die immer wieder als Wert erscheinen: -1 und -2. Der Wert -1 bedeutet sinngemäß "keine Angabe", d.h. der Befragte hat die Frage nicht beantwortet. Der Wert -2 erscheint, wenn die zugrundeliegende Frage nicht gestellt worden ist. Dies ist z.B. der Fall, wenn eine Person nicht verheiratet ist, dann macht eine Frage nach dem Heiratsjahr keinen Sinn und wird daher weggelassen.

Die Datei *pbiospe.dat* beinhaltet insgesamt 96768 Rekords. Wichtig ist die Betrachtung der Zustände: es werden 9 Zustände unterschieden; diese Anzahl ist niedriger als die Anzahl der aufgelisteten Zustände in *artkalen.dat*, in der 12 Zustände unterschieden werden. In der Tabelle 3.4 sind die Zustände und eine Häufigkeitsauszählung aufgelistet.

Tabelle 3.4: Ausprägungen der Zustände in *pbiospe.dat*

Zustand	Anzahl	Bedeutung
1	15410	Schule, Studium
2	15657	Lehre, Ausbildung
3	5214	Wehr-, Zivildienst
4	25571	Voll berufstätig
5	7736	Teilzeitbeschäftigt
6	7342	Arbeitslos
7	10213	Hausfrau, Hausmann
8	3870	im Ruhestand
9	5755	Andere Tätigkeit (Sonstiges)

3.3 Aufbereitung der Datensätze *artkalen.dat* und *pbiospe.dat*

Aus den Datensätzen *artkalen.dat* und *pbiospe.dat* selektieren wir jeweils die Spalten Personennummer, Zustand, Anfangszeitpunkt und Endzeitpunkt. Im weiteren Verlauf benötige ich zu jeder Person drei weitere Spalten: Geschlecht, Geburtsjahr und Stichprobenzugehörigkeit. Diese Spalten sind mit der Datei *ppfad.dat*⁷ innerhalb des SOEP gegeben. Zugleich empfiehlt sich die Reduktion der Daten, da der Umfang der Dateien die Schnelligkeit der Programmausführung bestimmt. In der vorliegenden Arbeit sollen die Personen der Stichproben A und B erhalten bleiben.

Für die Analysen des Projektes B6 sollten nur Personen herangezogen werden, die 1918 oder später geboren sind, da hier i. A. keine Lücken zwischen dem Biographieschema⁸ und dem Beginn der monatsbezogenen Daten des SOEP erwartet werden. Um einen allgemeineren Datensatz zu erstellen schien uns eine erweiterte Selektion nach dem Geburtsdatum ab 1900 angebracht⁹.

⁷Diese Datei ist im Umfang des SOEP enthalten. Sie besitzt für alle 27618 Personen, die jemals am SOEP teilgenommen haben, je einen Eintrag.

⁸Das Biographieschema erhebt die Daten vom 15. bis zum 65. Lebensjahr. Dies bedeutet für Personen, die 1918 geboren sind, daß ihre Daten der Erwerbskarriere durchgehend von 1933 bis 1983 vorliegen. 1983 beginnen die monatsbezogenen Daten durch die Erhebungen in den einzelnen Wellen.

⁹Dies ist auch nicht weiter problematisch, da vermutet werden kann, daß die Mehrheit der Personen ab dem 65. Lebensjahr in Rente gegangen ist, bzw. pensioniert ist.

3.3.1 Anpassung von *artkalen.dat*

Innerhalb der beiden Ursprungs-Datenfiles *pbiospe.dat* und *artkalen.dat* gibt es zwei offensichtliche Unterschiede, die im weiteren Verlauf der Spellerzeugung zu beheben sind: Dies ist die unterschiedliche Anzahl von Zuständen und die unterschiedliche Form der Darstellung der Anfangs- und Endzeitpunkte der Spells.

Die unterschiedlichen Zustandsräume beruhen auf einer Erweiterung der Zustände 1991 (8. Welle) und 1992 (9. Welle). In diesen Jahren wurden die Zustände Mutterschafts-Freistellung, nebenberufliche Tätigkeit und Kurzarbeit aufgenommen, die ihre Entsprechung in *artkalen.dat* gefunden haben. Zur Generierung eines allgemeingültigen Zustandsraumes muß eine Reduktion auf den Zustandsraum aus *pbiospe.dat* erfolgen: die Spalte Anfangszustand muß entsprechend des oben angegebenen Zustandsraumes neu codiert werden. An dieser Stelle erfolgt eine weitere Entscheidung: welchem Zustand sollen die neu hinzugekommenen Zustände Kurzarbeit, Mutterschafts-Freistellung und nebenberufliche Tätigkeit zugewiesen werden?

Kurzarbeit stellt eine von der Arbeitgeberseite her bestimmte verkürzte Dauer einer Vollzeitbeschäftigung dar, weshalb Kurzarbeit als Vollzeit-Erwerbstätigkeit betrachtet werden kann. Mutterschafts-Freistellung wurde bereits in den vorherigen Wellen unter dem Zustand "sonstiges" summiert, so daß hier in Analogie die Zuordnung zu "sonstiges" vorgenommen wurde. Die nebenberufliche Tätigkeit kann als Zusatztätigkeit betrachtet werden, so daß sie ebenfalls unter "sonstiges" plazierte werden kann.

Zur einfachen Handhabung der Start- und Endzeiten der Spelldaten und der Ereignismonate ist die Darstellung in Jahrhundertmonaten geeignet. Hierbei werden die Jahres- und Monatsangaben in eine neue Spalte überführt, indem sie folgendermaßen umgerechnet werden:

$$(\text{Jahresangabe}-1900)*12+\text{Monatsangabe}$$

Daraus ergeben sich z.B. folgende Werte:

$$\text{Oktober } 1959 = (1959-1900)*12+10 = 722$$

$$\text{Mai } 1993 = (1993-1900)*12+5 = 1121$$

In *artkalen.dat* liegen die Angaben folgendermaßen vor:

- 1 = Januar 1983
- 12 = Dezember 1983
- 13 = Januar 1984
- 24 = Dezember 1984
- etc., bis
- 132 = Dezember 1993

Die Umrechnung auf die Jahrhundertmonate ist somit:

$$\text{Jahrhundertmonat} = \text{Zahl} + (83 * 12) = \text{Zahl} + 996^{10}$$

Als eine kleine Erweiterung nehme ich die neue Spalte "Herkunft" in die Datei auf. Diese Spalte soll die Herkunft eines Spells festhalten. Die Werte dieser Spalte sind:

- 1 = Spell stammt aus *artkalen.dat*
- 2 = Spell stammt aus *pbiospe.dat*
- 3 = kombinierter Spell aus *artkalen.dat* und *pbiospe.dat*

3.3.2 Anpassung von *pbiospe.dat*

Innerhalb der Daten aus *pbiospe.dat* sind keine großen Veränderungen der Daten vorzunehmen: Die Spalte Spelltyp/Zustand ist durch Umkodierung der Werte mit den Ausprägungen in der Datei *artkalen.dat* in Einklang zu bringen (Siehe Tabelle 3.5).

Tabelle 3.5: Umkodierung der Werte in *pbiospe.dat*

<u>vorher</u>	<u>nachher</u>
1	6
2	3
3	7
4	1
5	2
6	4
7	8
8	5
9	9

Die bisher in Lebensjahren gegebenen Start- und Endzeitpunkte der Spells sind in Jahrhundertmonate umzurechnen. Als Übereinkommen im Projekt B6 wurde beschlossen, daß bei der Angabe von Start- und Endmonat des Spells beide Monate zum Spell gehören sollen. So ergibt sich folgende Umrechnung der Start- und Endzeitpunkte:

Startzeitpunkt: $(\text{Geburtsjahr}-1900 + \text{Altersangabe}) * 12 + 1$.

Endzeitpunkt: $(\text{Geburtsjahr}-1900 + \text{Altersangabe}) * 12 + 12$.

Die duration oder Dauer des Spells ergibt sich dann aus Endmonat-Anfangsmonat+1.

Eine Alternative besteht darin, den Endmonat als nicht zum Spell gehrend zu betrachten. Wodurch sich folgende Umrechnung ergibt:

Startzeitpunkt: $(\text{Geburtsjahr}-1900 + \text{Altersangabe}) * 12 + 1$.

Endzeitpunkt: $(\text{Geburtsjahr}-1900 + \text{Altersangabe}) * 12 + 13$.

¹⁰996 = Dezember 1982.

Die Duration oder Dauer des Spells berechnet sich dann aus Endmonat-Anfangsmonat.

3.4 Zwischenstand

Durch die genannten Schritte

- Erweiterung um die vier Grundvariablen aus *ppfad.dat*,
- Anpassung des Zustandsraumes der Spelltypen,
- Rekodierung der Anfangs- und Endzeiten der Spells,
- Erweiterung um die Spalte Herkunft des Spells.
- Eingrenzung der Daten auf Personen aus den Stichproben A und B,
- Eingrenzung der Daten auf Personen, die 1900 oder später geboren sind

und das anschließende Zusammenführen der Daten aus *artkalen.dat* und *pbios-pe.dat* ergibt sich die Datei *erwerb.dat*¹¹. In dieser Datei sind 125841 Rekords enthalten, 58617 von Männern und 67224 von Frauen. Durch die Selektion ist die älteste Person 1900 geboren, die jüngste 1979. Da es sich bei dem Geburtsdatum 1979 eine Fehlcodierung handelt, da Individualpersonen erst ab dem 16. Lebensjahr enthalten sein können, sollten diese Spells¹² aus dem Datensatz entfernt werden.

93493 Rekords sind von Personen aus der Stichprobe A und 32348 von Personen aus der Stichprobe B gegeben. Der Zustandsraum der Spells ist in der Tabelle 3.6 aufgelistet. Hier gibt es 1921 Spelle mit dem Wert 0 als Zustand. Diese müssen entfernt werden, da keine Zuordnung zu brauchbaren Werten vorgenommen werden kann.

Tabelle 3.6: Auszählung der Spalte Zustanin *erwerb.dat*

Wert	Anzahl	Bedeutung
0	1921	ehemaliges Label 99
1	34369	Vollzeit beschäftigt
2	11460	Teilzeit beschäftigt
3	14739	Lehre, Ausbildung
4	10856	Arbeitslos
5	6335	Rente, Ruhestand
6	16249	Schule, Hochschule
7	4481	Wehr-, Zivildienst
8	16895	Hausfrau, Hausmann
9	8536	Sonstiges

¹¹Zur Erstellung dieser Datei sind im Anhang A einige Informationen, wie die TDA-commandfiles und statistische Auszählungen, gegeben.

¹²Das sind nur zwei Spells einer Person, bei der offensichtlich das Geburtsdatum falsch ist. Daneben gibt es keine Spells von Personen mit dem Geburtsjahr 1978.

Der früheste Startzeitpunkt eines Spells ist 181 (Januar 1915), was auch mit dem frühesten Geburtsdatum (1900) in Einklang steht. Bis zum Startzeitpunkt 998 (Februar 1983¹³) liegt in jedem Jahr nur der Januar als Startzeitpunkt vor, da die retrospektive Erhebung jahresbezogen vorgenommen wurde. Erst mit Beginn der Erhebungswellen des SOEP ist jeder Monat als Startmonat der Spells möglich. Durch Ausfälle und Hinzukommen von Personen in das SOEP sind die Monate Januar aber weiterhin hoch besetzt. Da im Projekt B6 11 Wellen des SOEP vorliegen, setze ich den letztmöglichen Spellanfang auf Dezember 1994 (1140) fest. Die letzten 7 Spells liegen außerhalb dieses Zeitraumes und werden entfernt.

Als frühestes Spellende ist 192 (Dezember 1916) gegeben, was wiederum in Einklang mit den anderen Daten steht. Die Spellenden liegen bis 996 (Dezember 1982) ausschließlich im Dezember der Jahre und verteilen sich ab diesem Zeitpunkt auf das gesamte Jahr. Die Begrenzung des Zeitraumes liegt bei dem spätesten Ende eines Spells im Dezember 1994 (1140), wodurch 25 Spells entfernt werden müssen.

Der neu eingeführte Marker zeigt die Herkunft jeden Spells: 52036 Spells stammen aus *artkalen.dat* und 73805 aus *pbiospe.dat*.

Durch das Entfernen der angesprochenen Spells aus dem Datensatz verringert sich die Spellanzahl um 1947 auf 123894 Einträge.

3.5 Überschneidungen und Lücken

Eine Betrachtung der Anfangs- und Endzeiten der Spells zeigt, daß hier einige Probleme zu lösen sind. Die Spells liegen nicht schön aneinandergereiht vor, es gibt auch Spells, die sich überschneiden oder es liegen Lücken zwischen aufeinanderfolgenden Spells vor.

Vor der Behebung dieser Probleme sollen die Überschneidungen und Lücken genauer betrachtet werden¹⁴:

Überschneidungen

Überschneidung heißt: Der Startzeitpunkt eines Spells liegt vor dem Endzeitpunkt des vorhergehenden Spells einer Person. Eine einfache Kontrolle des Datensatzes auf diese Bedingung ergab über 60000 Überschneidungen, wobei zu beachten ist, daß die hohe Anzahl der Überschneidungen auf die Tatsache zurückzuführen ist, daß die Daten, die aus der Datei *pbiospe.dat* stammen, sich zwangs-

¹³D.h. der Februar 1983 ist der erste Monat neben dem Januar, in dem Spells beginnen können.

¹⁴Die Feststellung von Überschneidungen und Lücken ist nicht trivial. Der Aufwand zur Generierung von exakten statistischen Daten zu Überschneidungen steht in keinem Verhältnis zum Nutzen, weshalb hier darauf verzichtet wurde. Zur Analyse von Lücken mußte eine zusätzliche Datei erstellt werden, da leicht ersichtlich ist, daß auch bei geordneten Spells die Erkennung von Lücken aufgrund der unterschiedlichen Spelllängen nicht sofort möglich ist.

läufig mit den Daten aus der Datei *artkalen.dat* überschneiden müssen, da *pbios-pe.dat* die Zusammenführung aller Biographiedaten zum Erwerbsleben darstellt.

Überschneidungen sind nur dann ein Problem, wenn die Spelltypen nicht übereinstimmen. Für diesen Fall muß eine Rangfolge der Erwerbszustände erstellt werden. Der Spell mit der höheren Priorität überschreibt dann den Spell mit der niedrigeren Priorität. Die Prioritätenliste des Projektes B6 ist in der Tabelle 3.7 abgebildet.

Tabelle 3.7: Prioritätenliste der Zustände

Priorität	Spelltyp bzw. Zustand
1 = hoch	Vollzeiterwerbstätig
2	Teilzeiterwerbstätig
3	Betriebliche Ausbildung
4	Arbeitslos
5	Rente, pensioniert
6	Schule, Studium
7	Wehr-, Zivildienst
8	Hausfrau, -mann
9 = niedrig	sonstiges

Lücken

Bei Lücken gibt es zwei Möglichkeiten: sie können als Lücken erhalten bleiben oder aufgefüllt werden. Die Entscheidung, welche dieser beiden Alternativen angewandt wird, hängt auch von der Forschungsfrage ab. Da die Forschung des Projektes B6 auf den ganzen Lebenslauf ausgerichtet ist und andererseits Personen, deren Lebensläufe durch Nichtteilnahme am SOEP unterbrochen sind auch nicht ausgeschlossen werden sollten, fiel die Entscheidung, die Lücken aufzufüllen. Dies ist ein schwierigeres Problem, da zum Auffüllen der Lücken Annahmen über den zeitlichen Zwischenraum getroffen werden müssen.

Eine Auszählung der Lücken liefert folgendes Resultat:

Monate	Anzahl Spells
1 - 59	803
60 - 119	383
über 119	438

Bei Betrachtung des Fragebogens zur Biographieerhebung ist ein Grund für die Lücken bis zu 5 Jahren erkennbar, der für die Entscheidung, wie die Lücken geschlossen werden sollen, bedeutend ist: Das Biographieschema ist optisch in 5er Gruppen eingeteilt, wobei über jedem 5. Jahr eine große Markierung gegeben ist. Es ist leicht vorstellbar, daß Personen sich veranlaßt sahen, nur dies hervorgehobene Kästchen anzukreuzen statt der vorgesehenen 5. Aus dieser Berlegung heraus sollten Lücken mit einer Dauer bis zu 5 Jahren aufgefüllt werden, wenn die Zustände vor und nach dem Spell gleich sind. Bei ungleichen Zuständen kann

nicht entschieden werden, welcher Zustand in der Lücke vorherrscht, weshalb eine Aufspaltung der Lücken in zwei gleich große Teile angebracht erschien, die jeweils dem Spell vor bzw. nach der Lücke zugeordnet werden.

Zu den restlichen Lücken wurde entschieden, jene mit einer Dauer von weniger als 10 Jahren noch systematisch, die noch längeren nur ausnahmsweise aufzufüllen. Die Zustände "im Ruhestand", "arbeitslos" bzw. "in Wehr- oder Zivildienst" sollten dabei speziell behandelt werden, da der Zustand "im Ruhestand" zumeist der endgültige Austritt aus dem Arbeitsleben bedeutet, die Zustände "arbeitslos" bzw. "in Wehr- oder Zivildienst" nur von kurzer Dauer sind.

Aus diesen Betrachtungen heraus wurde schließlich folgende Behandlung der Lücken durchgeführt:

Sei SP = Typ des Spells vor der Lücke und SF = Typ des Spells nach der Lücke.

1. Lücken, deren Dauer kleiner 5 Jahren ist:

- ist $SP = SF$, so fülle die Lücke auf (es entsteht ein einziger Spell);
- ist $SP \neq SF$, so zerlege die Lücke in zwei gleiche Teile und füge diese Teile dem jeweiligen Spell zu; (bei ungerader Anzahl von Monaten wird dem vorderen Spell der überhängende Monat zugeteilt)

2. Lücken, deren Dauer größer oder gleich 5 Jahren und kleiner 10 Jahren ist:

- ist $SP = SF$, so fülle die Lücke auf;
- ist $SP = 5$ (im Ruhestand), so füge die Lücke an den vorderen Spell an;
- ist $SP = 4$ (arbeitslos) oder 7 (Wehr- oder Zivildienst), so füge ein Jahr der Lücke an den vorderen Spell an, den Rest zum folgenden Spell;
- ist $SF = 4$ (arbeitslos) oder 7 (Wehr- oder Zivildienst), so füge ein Jahr der Lücke dem hinteren Spell zu, den Rest der Lücke zum vorderen Spell;
- ist $SP \neq SF$ und weder SF noch SP entsprechen den Zuständen 4 oder 7, so teile die Lücke in zwei gleiche Teile und füge diese Teile dem jeweiligen Spell zu;

3. Lücken, deren Dauer größer oder gleich 10 Jahren ist:

- ist $SP = 5$, so füge die Lücke dem vorderen Spell an;
- alle anderen Lücken bleiben erhalten;

Mit den Entscheidungen, wie mit Überschneidungen und Lücken umzugehen ist, kann nun die Bearbeitung der Erwerbsspell fortgesetzt werden. Für diese Arbeit

wurde ein C-Programm geschrieben, welches die Datei *erwerb.dat* einliest und die Erwerbsspell aufarbeitet. Als Ergebnis liegt die Datei *erwerb_spells.dat* vor.

Wie das C-Programm die Arbeit im Einzelnen ausführt, ist nicht weiter interessant und soll daher hier nicht erläutert werden. Entscheidend ist nur, daß die Überschneidungen und Lücken anhand der oben dargestellten Lösungen überarbeitet wurden. Darüber hinaus generiert das Programm zusätzliche Spalten:

- Endzustand:

In dieser Spalte wird der Zustand/Spelltyp des folgenden Spells eingetragen, so daß erkennbar wird, welcher Zustand nach dem vorliegenden Spell folgt. Ist der vorliegende Spell der letzte Spell der Person oder folgt eine Lücke, so ist Anfangszustand = Endzustand. An dieser Kombination ist damit auch eine Rechtszensierung zu erkennen.

- Zensierung:

Diese Spalte hat die in der Tabelle 3.8 aufgelisteten Werte

Tabelle 3.8: Werte und Label zur Zensierung

Wert	Beschreibung
0	Keine Zensierung
1	Der Spell ist links zensiert. Dies ist immer bei dem ersten Spell einer Person der Fall und bei dem ersten Spell nach einer verbleibenden Lücke.
2	Der Spell ist rechts zensiert. Dies ist immer bei dem letzten Spell einer Person der Fall und bei einem Spell vor einer Lücke.
3	Der Spell ist rechts- und linkszensiert.

- Marker zur Lückenbehandlung:

In dieser Spalte wird vermerkt, ob und wie eine Lücke behandelt worden ist. Für die getroffene Entscheidung ergeben sich die in der Tabelle 3.9 aufgelisteten Möglichkeiten.

Tabelle 3.9: Werte und Label zur Lückenbehandlung

Werte	Beschreibung
0	keine Lückenbehandlung;
1	eine Lücke innerhalb des Spells wurde aufgefüllt;
2	der linke Teil einer Lücke wurde rechts an diesen Spell angefügt;
4	der rechte Teil einer Lücke wurde links an diesen Spell angefügt;
10	die ganze Lücke wurde rechts an den Spell angefügt (Anfangszustand war Rente);
20	die ersten 12 Monate einer Lücke wurde rechts an den Spell angefügt (Anfangszustand ist Wehr-, Zivildienst oder arbeitslos);
40	der Rest der Lücke (vom Wert 20) wurde links an den Spell angefügt;
100	die letzten 12 Monate einer Lücke wurden links an den Spell angefügt (Endzustand ist Wehr-, Zivildienst oder arbeitslos);
200	der Rest der Lücke (vom Wert 100) wurde rechts an den Spell angefügt;

Zusätzlich zu den in der Tabelle 3.9 dargestellten Werten ist noch jede Kombination möglich, da jeder Spell mehr als eine Lücke oder Lückenteil aufgenommen haben kann: z.B. wurde eine Lücke innerhalb des Spells aufgefüllt (1) und gleichzeitig der linke Teil einer Lücke rechts angefügt (2), so ergibt sich der Wert 3¹⁵.

- Marker zur Herkunft des Spells:

Wert	Beschreibung
1	Der Spell stammt rein aus dem Kalendarium
2	Der Spell stammt rein aus dem Biographieschema
3	Der Spell ist aus Kalendarium und Biographieschema zusammengesetzt.

Durch die Bearbeitung entsteht der Datensatz *erwerb_spells.dat*, der im folgenden Abschnitt näher erläutert wird.

3.6 Einige Informationen zu *erwerb_spells.dat*

Die Datei *erwerb_spells.dat* stellt für Personen der Stichproben A und B, die 1900 oder später geboren sind, die Daten ihrer Erwerbskarriere in Spelldatenformat zur Verfügung. Der Aufbau der Datei ist in der Tabelle 3.10 dargestellt.

¹⁵Nicht jeder Wert oder jede Kombination tritt später auf, es handelt sich hierbei um die möglichen Werte.

Tabelle 3.10: Aufbau der Datei *erwerb_spells.dat*

Spalte	Beschreibung
1	Personennummer
2	Geschlecht
3	Geburtsdatum
4	Stichprobenzugehörigkeit
5	Spellnummer
6	Anfangszustand
7	Endzustand
8	Anfangsmonat
9	Endmonat
10	Zensierung
11	Dauer
12	Marker für die Herkunft
13	Marker für die Lückenbehandlung

Welche Werte im einzelnen möglich sind, ist in den vorherigen Abschnitten dargestellt worden oder aus Anhang B ersichtbar.

Zu den statistischen Daten¹⁶ der Datei *erwerb_spells.dat*:

- Insgesamt gibt es 63364 Spells.
- 29359 Spells sind von Männern und 34005 von Frauen.
- Die älteste Person ist 1900 geboren, die jüngste 1977.
- 47873 Spells stammen von Personen aus der Stichprobe A und 15491 von Personen aus der Stichprobe B.
- Die maximale Spellanzahl einer Person ist 25.
- Der Zustandsraum von Anfangs- und Endzustand (Origin- und Destinationstate) umfaßt die Zustände 1 bis 9.
- Der früheste Startzeitpunkt eines Spells ist 181 (Januar 1915), der späteste 1129 (Januar 1994). Der Anfang von Spells die vor der monatsbezogenen Erhebung des SOEP (997 = Januar 1983¹⁷) liegen, fangen erwartungsgemäß vor allem im Januar des Jahres an. Ausnahmen, die durch die Lückenauffüllung einer ungeraden Anzahl von Jahren entstanden sind, starten immer im Juli des Jahres.

¹⁶Siehe auch Anhang B.

¹⁷Die Daten der Erwerbskarriere werden auch bezogen auf das Vorjahr erhoben. Somit ergeben sich mit der ersten Befragung 1984 die Daten aus dem Jahr 1983.

- Der früheste Endzeitpunkt eines Spells ist 192 (Dezember 1915), der späteste 1140 (Dezember 1994). Das Spellende von Spells die vor der monatsbezogenen Erhebung des SOEP (1997) liegen, enden erwartungsgemäß vor allem im Dezember des Jahres. Ausnahmen, die durch die Lückenauffüllung einer ungeraden Anzahl von Jahren entstanden sind, enden immer im Juni des Jahres.

Die bei weitem höchste Anzahl von Spells endet im Monat 1128 (Dezember 1993): es sind 8848 Spells. Dies ist gegeben durch die Befragung der 11. Welle 1994, in der die Erwerbsdaten des vorherigen Jahres erfragt wurden. Problematischer sind die 189 Spells, die den Monat 1140 (Dezember 1994) als Endzeitpunkt haben. 185 dieser Fälle ergeben sich durch die Befragung der Biographiedaten, die für diese Personen im Jahr 1994 stattgefunden hat. Dadurch ergibt sich der Monat 1140 (Dezember 1994) als Endmonat der Spells durch die Behandlung dieser jahresbezogenen Daten. Es bleiben 4 Fälle¹⁸ von den 189 Spells übrig, die sich durch falsche Daten in der Datei *pbiospe.dat* ergeben.

- Die Zensierung zeigt 31930 unzensierte Spells, 14877 links-, 14877 rechts- und 1630 beidseitig zensierte Spells.
- Die kürzeste Dauer eines Spells ist 1 Monat, die längste 768 Monate (64 Jahre). Erwartungsgemäß sind die Monate, die exakten Jahresangaben entsprechen, überrepräsentiert.
- Der Marker für die Herkunft¹⁹ der Spells zeigt auf, daß 5503 Spells rein aus den Biographiedaten stammen, 53282 Spells rein aus dem 11-Jahres-Fenster stammen und 4579 Spells aus beiden Datenquellen zusammengesetzt wurden.
- Mit dem Marker für die Lückenbehandlung ist erkennbar, daß:
 - 60511 Spells nicht verändert wurden (Wert=0),
 - bei 1054 Spells mindestens eine Lücke gefüllt wurde (Wert=1),
 - bei 858 der linke Teil einer Lücke rechts angehängt wurde (Wert=2),
 - bei 10 Spells sowohl eine Lücke gefüllt als auch der linke Teil einer Lücke rechts angehängt worden ist (Wert=3),
 - bei 851 Spells der rechte Teil einer Lücke links angehängt wurde (Wert=4),
 - bei 17 Spells mindestens eine Lücke gefüllt und der rechte Teil einer Lücke links angehängt wurde (Wert=5),

¹⁸Das sind die Personen mit den Nummern 304801, 436305, 512402 und 521902.

¹⁹Dabei darf nicht vergessen werden, daß die Daten aus dem 11-Jahres-Fenster immer die höhere Priorität bei der Übernahme besitzen. So ergibt sich die relativ geringe Anzahl der Spells rein aus den Biographiedaten.

- bei 2 Spells eine Lücke rechts angehängt wurde (Wert=10),
- bei 6 Spells die ersten 12 Monate einer Lücke rechts angehängt wurde (Wert=20),
- bei 6 Spells der Rest der Lücke links angehängt wurde (Wert=40),
- bei 1 Spell mindestens eine Lücke gefüllt und der Rest einer Lücke links angehängt wurde (Wert=41),
- bei 4 Spells die letzten 12 Monate einer Lücke links angehängt wurde (Wert=100),
- bei 1 Spell die ersten 12 Monate einer Lücke rechts und die letzten 12 Monate einer Lücke links angehängt wurde (Wert=120),
- bei 5 Spells der Rest einer Lücke rechts angehängt wurde (Wert=200).

Kapitel 4

Daten zu Ehen und Kindern

Als wichtige Statuspassagen des Lebenslaufs betrachtet das Projekt B6 die Eheschließung, Ehescheidung und die Geburt von Kindern. Diese Ereignisse werden in einer Datei zusammengefaßt, wodurch eine leichte Verbindung mit den Daten über das Erwerbsleben möglich ist¹.

In diesem Kapitel gebe ich zuerst eine Übersicht über die Anforderungen des Projektes B6 an den Ergebnisdatensatz. Zwei Abschnitte befassen sich dann getrennt mit der Erzeugung der Daten von Ehen bzw. Kindern. In diesen Abschnitten erfolgt zunächst eine Beschreibung der Ausgangsdatensätze des SOEP. Daran schließt sich die Darstellung der vorgenommenen Veränderungen zur Erzeugung des Ergebnisdatensatzes an.

4.1 Anforderungen an den Datensatz *ereignisse.dat*

Die Daten zu Ehen und Kindern sollen in die Datei *ereignisse.dat* übernommen werden. Die Grundvariablen Geschlecht, Geburtsjahr und Stichprobenzugehörigkeit sollen ebenfalls in diesen Datensatz Eingang finden, so daß folgende Daten mit dem Datensatz *ereignisse.dat* bereitgestellt werden:

- Personennummer
- Geschlecht
- Geburtsdatum
- Stichprobenzugehörigkeit
- Heiratsdatum als Anfang einer Ehe
- Datum von Scheidung oder Tod des Partners als Ende einer Ehe

¹Z.B. über die Splittung der Erwerbssphäre zu den Zeitpunkten des Ereigniseintrittes.

- Kindergeburten

Die Datei *ereignisse.dat* soll den in der Tabelle 4.1 ausgeführten Aufbau vorweisen.

Tabelle 4.1: Aufbau der Datei *ereignisse.dat*

Spalte	
1	Personennummer
2	Geschlecht
3	Geburtsjahr
4	Stichprobenzugehörigkeit
5	Jahr des Ereigniseintrittes
6	Monat des Ereigniseintrittes
7	Ereignisart
8 - 10	Zusatzinformationen zum Ereignis

Kurze Beschreibung der Spalten:

Spalte 1: Personennummer

Die eindeutige Personennummer;

Spalte 2: Geschlecht

Die Angabe des Geschlechts einer Person: 1=männlich, 2=weiblich.

Spalte 3: Geburtsjahr

Das Geburtsjahr der Person, angegeben als Jahreszahl.

Spalte 4: Stichprobenzugehörigkeit

Angabe der Stichprobenzugehörigkeit der Person: 1 = Stichprobe A, 2 = Stichprobe B.

Spalte 5: Jahr des Ereigniseintrittes

Die Angabe des Jahres, in dem das Ereignis eingetreten ist, als Jahreszahl.

Spalte 6: Monat des Ereigniseintrittes

Von Ereignissen, die innerhalb des 11-Jahres-Fensters eingetreten sind, sind die Monatsangaben bekannt und können hier übernommen werden. Retrospektiv erhobene Ereignisse weisen diese Angabe nicht auf, sie haben als Wert eine 0. Bei späteren Analysen bleibt es somit dem Forscher überlassen, ob er jahresbezogene oder monatsbezogene Daten bevorzugt. Bei einer Entscheidung zugunsten monatsbezogener Daten, ist zu überlegen, welcher Monat den Ereignissen mit dem Wert 0 zugewiesen werden soll.

Spalte 7: Ereignisart

Da unterschiedliche Ereignisse in diese Datei übernommen werden sollen, sind diese anhand von Kennzahlen zu unterscheiden. Ich habe Ereignisse in Gruppen zusammengefaßt, so daß ein Spektrum an Kennzahlen für jede Gruppe zur

Verfügung steht. Bei einer späteren Erweiterung der Ereignisse ist damit weiterhin für eine sinnvolle Zuordnung von Kennzahlen zu den Ereignissen gesorgt. Meine Zuordnung sieht folgendermaßen aus:

Bereich 100 bis 199	Ehe
Bereich 200 bis 299	sonstige Partnerschaft ²
Bereich 300 bis 399	Kinder
Bereich 400 bis 499	Bildung
199	Problemfall ³

Im Detail bedeutet das:

Wert	Bedeutung
110	Anfang einer Ehe
120	Ende einer Ehe
300	Geburt eines Kindes

Spalten 8 bis 10:

Viele Ereignisse erfordern zusätzliche Informationen. Für diesen Zweck werden drei Spalten zur Verfügung gestellt: 8 bis 10. Diese drei Spalten stehen also in enger Verbindung mit dem in der Spalte 7 genannten Ereignis.

Für das Ereignis 110 (Anfang einer Ehe) sind folgende Angaben⁴ vorgesehen:

Spalte 8:	1	retrospektiv angegebenes Anfangsjahr einer Ehe
	6	Anfang einer Ehe durch Anfang der Befragung (Familienstand)
	8	Anfang einer Ehe durch einen angegebenen Heiratsmonat
	12	Anfang einer Ehe durch Wechsel der Partnernummer
	14	Anfang einer Ehe durch Wechsel des Partnerzeigers
	21	Zugewiesene Heirat an diese Person als Korrektur der falschen Daten
Spalte 9:		Angabe der Partnernummer; 0 falls der Partner unbekannt ist
Spalte 10:		Nummer der Ehe

Für das Ereignis 120 (Ende einer Ehe):

²Im vorliegenden Arbeitspapier wurden Partnerschaften ohne Trauschein nicht behandelt. Da eine spätere Erweiterung nicht auszuschließen ist, wurde dieser Gruppe bereits ein Bereich von Kennzahlen zugeordnet.

³Eine Person, der wegen falscher oder unvollständiger Daten ein Ereignis nicht eindeutig zugewiesen werden kann. Auf diese Fälle wird in den Abschnitten der Datenerzeugung gesondert eingegangen.

⁴Die scheinbar nicht sortierte Nummerierung faßt gleichartige Ereignisse zu Gruppen zusammen, um eine spätere Selektion zu erleichtern.

Spalte 8:	2	retrospektiv angegebenes Ende einer Ehe durch Scheidung
	3	retrospektiv angegebenes Ende einer Ehe durch den Tod des Partners
	4	retrospektiv angegebenes Ende einer Ehe ohne Angabe des Grundes
	9	Ende einer Ehe durch einen angegebenen Scheidungsmonat
	10	Ende einer Ehe durch den angegebenen Tod des Partners
	13	Ende einer Ehe durch einen Wechsel in der Partnernummer
	15	Ende einer Ehe durch einen Wechsel des Partnerzeigers
	22	Ende einer Ehe durch die Korrektur der Daten: Eine Scheidung wurde eingefügt
	23	Ende einer Ehe durch die Korrektur der Daten: Der Tod des Partners wurde eingefügt
	24	Ende einer Ehe durch die Korrektur der Daten: Grund des Eheendes ist unbekannt
Spalte 9:		Angabe der Partnernummer; 0 falls der Partner unbekannt ist
Spalte 10:		Nummer der Ehe

Die Werte 1 bis 4 beziehen sich auf die retrospektiv erhobenen Ehedaten. Der Wert 6 sollte vergeben werden, wenn eine bereits verheiratete Person in das SOEP hineinkommt, die jedoch nicht retrospektiv befragt wurde und von der die vollständigen Daten nicht vorliegen. Die Werte 8 bis 10 gehören zu Ehen, die im 11-Jahres-Fenster geschlossen oder beendet wurden. Ich habe noch einige vermutete Problemfälle in Betracht gezogen, denen die Werte 12 bis 15 zugeordnet werden sollten: Personen, denen auf Grund von Veränderungen in der Partnernummer oder des Familienstandes der Anfang oder das Ende einer Ehe hinzugefügt werden sollte. Die Werte 21 bis 24 schließlich sollten Personen erhalten, bei denen die Daten individuell mit Durchsicht der Datensätze zu ändern waren. Die Werte 6 und 12-15 mußten später nicht vergeben werden.

Für das Ereignis 300 (Geburt eines Kindes) sind bisher keine weiteren Angaben in den Spalten 8 bis 9 geplant. Mögliche Erweiterungen könnten sich für die Spalte 7 (Ereignisart) durch die Einführung der Altersübergänge der Kinder ergeben: z.B. 306: ein Kind erreicht das sechste Lebensjahr. Die Spalte 10 gibt die Anzahl der Kinder dieser Person in diesem Jahr an.

Wie bereits bei den Erwerbspells beschrieben, sollen auch bei den jetzt zu erstellenden Datensätze die Selektionen von Personen der Stichproben A und B und ab dem Geburtsjahr 1900 erfolgen.

4.2 Partnerschaften, Partner, Partnerzeiger

Die Erkennung von Partnerschaften stellt innerhalb der Datenaufbereitung des SOEP eine besondere Herausforderung dar. Dies ist vor allem der Fall, wenn für Analysen die zugehörigen Partner identifiziert werden sollen.

Da bei der retrospektiven Biographieerhebung der Familiendaten nur Ehen erhoben wurden, sind nicht eheliche Partnerschaften nur innerhalb des 11-Jahres-

Fensters zu rekonstruieren. Bei den Analysen des Projektes B6 beschäftigen sich die Fragestellungen mit Ehen, so daß Partnerschaften ohne Trauschein nicht bearbeitet wurden. Die spätere Erweiterung ist durch die allgemeine Struktur der Datei *ereignisse.dat* jedoch einfach möglich.

Im weiteren Verlauf konzentriere ich mich also auf die Erkennung von Ehen und Ehepartnern. Als Ehedaten sollen Anfang und Ende einer Ehe sowie der Ehepartner rekonstruiert werden. Trennungszeiten der Ehepartner vor der Scheidung bleiben unbeachtet.

Ehedaten liegen in zwei Quellen vor. Erstens wurden innerhalb der Biographiedatenerhebung die bis zum Erhebungszeitpunkt stattgefundenen Ehen retrospektiv auf Jahresbasis erfragt (Datensatz *biolela.dat*). Zweitens stehen im 11-Jahres-Fenster Heirats-, Scheidungsmonate und Monate des Todes eines Partners zur Verfügung (Datensätze *ap.dat* bis *ak.dat*).

Die Erkennung des Partners ist nur im 11-Jahres-Fenster möglich; nur hier kann die Zuordnung einer Partnernummer über den Partnerzeiger erfolgen. Der Partnerzeiger bestimmt dabei, wie die Partnernummer dieses Jahres zu interpretieren ist:

- 1 = die Partnernummer ist der eindeutige Ehepartner
- 2 = die Partnernummer ist der eindeutige Lebenspartner
- 3 = die Partnernummer ist der vermutliche Ehepartner
- 4 = die Partnernummer ist der vermutliche Lebenspartner
- 9 = Partner ja, aber wer?; keine Angabe in der Partnernummer

In den folgenden Abschnitten soll nacheinander auf die Aufbereitung der Datensätze *biolela.dat* und der Daten aus dem 11-Jahres-Fenster eingegangen werden.

4.2.1 Ehedaten aus *biolela.dat*

Grundlage der folgenden Arbeiten stellt der Datensatz *biolela.dat* dar. Hier wurden retrospektiv erhobene Daten der Personen zeilenweise aufgenommen, so daß jeder Person maximal ein Rekord zugeordnet werden kann.

Der Datensatz *biolela.dat* beinhaltet für jede Person 262 Spalten. Nach einer ersten Überprüfung des Datensatzes, stellte sich heraus, daß einige Daten nur unzureichend enthalten sind. So stehen von insgesamt 21747 Personen nur 1143 Geburtsdaten in der Spalte 10 bereit. Daher besteht einer der ersten Schritte darin, die erwähnten vier Grundvariablen zu vervollkommen um die erforderlichen Selektionen durchzuführen.

Für die Erstellung der Ehedaten sind die Spalten⁵

- 46 (Beginn 1. Ehe), 47 (Beginn 2. Ehe), 48 (Beginn 3. Ehe),

⁵Der Datensatz *biolela.dat* bietet sehr viele Informationen. Ich beschränke mich hier auf die für das Anliegen wichtigen Spalten. Weitere Informationen zum Aufbau der Datei findet sich in der Datei *biolela.var*, die mit der Datei *biolela.dat* geliefert wird.

- 49 (Status 1. Ehe), 50 (Status 2. Ehe), 51 (Status 3. Ehe),
- 52 (Ende 1. Ehe), 53 (Ende 2. Ehe) und 54 (Ende 3. Ehe)

entscheidend⁶.

Eine statistische Betrachtung der Spalten 46 bis 54⁷ ergibt folgendes Bild:
(in den Tabellenfeldern steht die Anzahl, ein Leerfeld bedeutet 0)

Wert	Var. 46 Anfang 1. Ehe	Var. 47 2. Ehe	Var. 48 3. Ehe	Var. 52 Ende 1. Ehe	Var. 53 2. Ehe	Var. 54 3. Ehe
-2	4623	15912	17451	15326	17546	17723
-1	4095	4007	4006	4061	4009	4017
1		672			8	
2		174			2	
3		73	1		1	
> 12	13029	909	289	2360	181	7

Die Werte -2 und -1 wurden oben bereits erläutert. Die anderen Werte geben das Jahr des Ereignisses an, bzw. sollen das Jahr angeben. Aus dieser Auflistung läßt sich jedoch ein krasser Fehler der Daten erkennen, denn die Jahreszahlen 1 (für 1901), 2 (für 1902) und 3 (für 1903) können als Heiratsjahr nicht in dieser Häufigkeit auftauchen, da nur 277 Personen überhaupt vor 1901 geboren sind und bei einem Heiratsalter von mindestens 10 Lebensjahren bleiben davon nur 10 Personen übrig. Der Grund dieses Fehlers liegt in der Reihenfolge der Spalten, wie sie für die in der zweiten Welle erhobenen Biographiedaten erstmals benutzt worden ist. Hier liegt eine etwas andere Reihenfolge der Spalten für die retrospektiv erhobenen Daten der Ehen zugrunde:

- Beginn 1. Ehe, Status 1. Ehe, Ende 1. Ehe,
- Beginn 2. Ehe, Status 2. Ehe, Ende 2. Ehe,
- Beginn 3. Ehe, Status 3. Ehe und Ende 3. Ehe.

In dieser Reihenfolge ist offensichtlich ein Teil der in *biolola.dat* enthaltenen Daten gegeben. Dies ist bei der späteren Datenaufbereitung zu beachten.

Erstaunlich und unerklärlich bei der Betrachtung der Daten ist die Tatsache, daß die zwischen den Spalten 48 und 52 liegenden Spalte des Status' diese Struktur nicht aufweisen, d.h. hier sind keine Jahresangaben, die auf einen Eheanfang oder -ende schließen ließen, zu sehen:

⁶Es wurden nur drei Ehen retrospektiv erhoben. Dies ist jedoch nicht problematisch, da, wie sich später zeigen wird, nur sehr wenige Personen die dritte Ehe als beendet angeben und somit Kandidaten für eine vierte weitere Ehe wären.

⁷Eine ausführlichere Liste befindet sich im Anhang C.

Wert	Spalte 49 Status 1. Ehe	Spalte 50 Status 2. Ehe	Spalte 51 Status 3. Ehe
-2	5449	16742	17670
-1	4312	4033	4034
1	9607	766	39
2	1270	120	3
3	1109	86	1

Die Spalte für den Status hat neben den Werten -2 und -1 die Ausprägungen 1 (Ehe besteht zum Befragungszeitpunkt), 2 (Ehe ist durch eine Scheidung beendet worden) und 3 (Ehe wurde durch den Tod des Partners beendet). Darüber hinaus gibt es keine Angaben.

Im ersten Schritt erweitere ich die aus *biolela.dat* selektierten Daten zu den Ehedaten um die drei Grundvariablen Geschlecht, Geburtsjahr und Stichprobenzugehörigkeit, die der Datei *ppfad.dat* entnommen werden. Dadurch sind die Selektionen bezüglich Stichprobenzugehörigkeit (Stichprobe A und B) und Geburtsjahr (1900 und später) durchführbar.

Aus diesen Daten können jetzt nicht einfach Ereignisdaten erstellt werden, da Überschneidungen mit Eheangaben aus dem 11-Jahres-Fenster möglich sind. Für jede Person müssen also zuerst zusätzlich alle Angaben bezüglich Ehen (Heirat, Scheidung, Tod des Partners, Partnernummer, Partnerzeiger) aus den Daten des 11-Jahres-Fensters gewonnen und in eine gemeinsame Datei übertragen werden, aus der dann die Ehedaten rekonstruiert werden.

Einige statistische Informationen zu den retrospektiv erhobenen Ehedaten

Bei Betrachtung der statistischen Daten zu den retrospektiv erhobenen Ehedaten⁸ ist erkennbar:

- Die Auszählung der Spalte ANFANG1EHE (Anfang der ersten Ehe) zeigt, daß von insgesamt 6872 Personen keine Angaben zu einer ersten Ehe vorliegen. Das sollten neben den Unverheirateten vor allem Kinder sein.
- Die Spalte ANFANG2EHE (Anfang der zweiten Ehe) zeigt den oben beschriebenen Fehler: bei 841 Personen ist als Anfang der zweiten Ehe 1, 2 oder 3 eingetragen. Da dies in der Mehrzahl dieser Fälle ein Fehler sein sollte, werden im weiteren Verlauf die Daten dieser Personen anders aufbereitet, für sie werden die Daten als in der alten Reihenfolge vorliegend behandelt.

Desweiteren ist für 14780 Personen kein Anfang einer 2. Ehe angegeben. Der letzte Anfang einer zweiten Ehe ist 1994.

⁸Siehe auch Anhang C.

- Der Anfang einer dritten Ehe (ANFANG3EHE) ist bei 15891 Personen nicht gegeben. Eine Person hat als Anfangsjahr der Ehe 1903 angegeben. Hierbei handelt es sich um einen Fehler, da diese Person 1909 geboren ist und als Anfangsjahr der ersten Ehe 1934 angegeben hat. Die Angabe zum Beginn der dritten Ehe bleibt bei der Datenaufbereitung unbeachtet. Die letzten dritten Ehen beginnen 1993.
- Die drei Spalten des Status (STATUS1EHE, STATUS2EHE, STATUS3EHE) weisen nur die erwarteten Werte -2, -1, 1, 2 und 3 auf. Von den dritten Ehen werden 2 als beendet angegeben⁹. Nur diese Personen sind Kandidaten für eine mögliche vierte Ehe, die aber nicht retrospektiv erhoben wurde¹⁰. Wegen der geringen Anzahl der möglichen vierten Ehen, sollte kein relevanter Fehler durch eine fehlende Ehe entstehen.
- Die Spalte ENDE1EHE (Ende der ersten Ehe) zeigt keine Besonderheiten.
- Bei der Spalte ENDE2EHE (Ende der zweiten Ehe) erscheint wiederum der oben erwähnte Datenfehler: die Angaben 1, 2 und 3 sind 11 Mal gegeben und sind als Fehler anzusehen. Bei diesen Personen werden die Daten als in der alten Reihenfolge vorliegend betrachtet und aufbereitet.
- Die Spalte ENDE3EHE (Ende der dritten Ehe) zeigt Jahresangaben von 5 dritten Ehen, die beendet wurden.
- Es sind 7947 Männer und 8184 Frauen im Datensatz gegeben.
- In diesem Datensatz ist auch die Person mit dem Geburtsjahr 1979 wieder vorhanden und muß daher später entfernt werden.
- Es sind 11603 Personen aus der Stichprobe A und 4528 Personen aus der Stichprobe B gegeben.

4.2.2 Ehedaten aus dem 11-Jahres-Fenster

Für die Gewinnung von Ehedaten aus dem 11-Jahres-Fenster sollen unterschiedliche Angaben herangezogen werden. Durch die jährlich stattfindenden Befragungen stehen mit Bezug zur Ehe folgende Daten monatsbezogen zur Verfügung:

- Heirat,
- Scheidung und
- Tod des Partners.

⁹Diese Zahl vergrößert sich auf maximal 7, wenn die Spalte ENDE3EHE betrachtet wird.

¹⁰Es ist jedoch jederzeit möglich, durch eine weitere Ehe innerhalb der monatsbezogenen Datenerhebung der 11 Wellen auf vier oder mehr Ehen zu kommen.

Hinzugefügt werden jahresbezogen nach den erhobenen Angaben:

- Familienstand,
- Partnernummer und
- Partnerzeiger.

Bei der Selektion der Daten zu Hochzeit, Scheidung und Tod des Partners ist zu beachten, daß diese Daten auf zwei Spalten¹¹ verteilt für jedes Jahr vorliegen. Dies liegt an dem Zeitpunkt der Befragung, der nicht am Ende, sondern innerhalb des Jahres liegt. Für die Datei *familie.dat*, die die Familiendaten aus dem 11-Jahres-Fenster aufnehmen soll, müssen beide Spalten in eine überführt werden. Hier ist zu beachten, daß z.B. je nach Heiratsmonat, der Monat in einer oder beiden Spalten enthalten ist. An dieser Stelle können auch Inkonsistenzen auftauchen, falls beide Monatsangaben nicht übereinstimmen.

Der einfache Weg, um eine gemeinsame Spalte zu erstellen, ist es, die erste Angabe¹² zu übernehmen, und nur wenn dort keine Angabe ist, die zweite Spalte heranzuziehen.

Die Angaben müssen zusammen mit der Personennummer aus den jährlichen Datenquellen *ap.dat*, *bp.dat* bis *kp.dat* in eine Datei selektiert werden. An dieser Stelle bietet es sich an, die Angaben zu den Geburten von Kindern mit zu selektieren, da diese im weiteren Verlauf benötigt werden. Bei der Selektion der Daten kann auch sofort die Eingrenzung erfolgen, die für alle bisher erstellten Dateien vorgenommen wurden: es sollen nur Personen aus den Stichproben A oder B und einem Geburtsjahr ab 1900 aufgenommen werden. Der Aufbau der dadurch entstandenen Datei *familie.dat* ist in der Tabelle 4.2 erläutert.

Tabelle 4.2: Aufbau der Datei *familie.dat*

Spalte	Bedeutung
1	Personennummer
2 - 12	Partnerzeiger 84 - 94
13 - 23	Partnernummer 84 - 94
24 - 34	Familienstand 84 - 94
35 - 45	Heiratsmonat 84 - 94
46 - 56	Scheidungsmonat 84 - 94
57 - 67	Todesmonat des Partners 84 - 94

¹¹Die Befragung jeder Welle findet irgendwann im Jahr (zumeist im März) statt. Es werden die Veränderungen des Vorjahres erhoben (erste Spalte) und die Veränderung des gegenwärtigen Jahres bis zum Erhebungsmonat (zweite Spalte).

¹²In der Annahme, daß die Erinnerung der Ehegatten an die Heirat aktueller ist, da diese erst vor ca. einem halben Jahr war.

Einige statistische Informationen zu *familie.dat*

Die Tabelle 4.3 liefert einen Überblick über die Anzahl der Ereignisse Heirat, Scheidung und Tod des Partners von Personen der Stichproben A und B (siehe Seite 4), die 1900 oder später geboren sind (Siehe auch Anhang C).

Tabelle 4.3: Auszählung der Heiraten, Scheidungen und Tode des Partners aus dem 11-Jahres-Fenster

Jahr	Heirat	Scheidung	Tod des Partners
1984	175	36	33
1985	216	30	34
1986	268	41	46
1987	214	35	41
1988	217	35	36
1989	230	34	37
1990	199	31	36
1991	254	45	32
1992	215	27	39
1993	237	48	42
1994	31	17	3

4.2.3 Zusammenstellung der Ehe-Ereignisse

Die Daten der Dateien *biolala.dat*, *ppfad.dat* und *familie.dat* können jetzt zusammengebracht werden, wobei sich die Zusammenführung an der Personennummer orientiert, so daß für jede Person weiterhin genau ein Rekord gegeben ist (Siehe dazu auch Anhang C).

Die so erstellte Datei *ehen.dat* muß nun überarbeitet werden. Herauszufiltern sind:

- die Anfänge von Ehen, die gegeben sind durch
 - die retrospektive Angabe eines Eheanfangs,
 - einen Hochzeitsmonat im 11-Jahres-Fenster,
 - einen Wechsel im Familienstand,
 - einen Wechsel im Partnerzeiger oder
 - das Hereinkommen ins SOEP mit dem Familienstand "verheiratet".
- die Endpunkte von Ehen, die gegeben sind durch
 - die retrospektive Angabe einer Scheidung (über Status der Ehe)
 - die retrospektive Angabe des Todes des Partners
 - einen Scheidungsmonat im 11-Jahres-Fenster,
 - einen Todesmonat des Partners

- den Wechsel im Familienstand oder
- den Wechsel im Partnerzeiger.

Ein Wechsel in Familienstand oder Partnerzeiger, der ohne Angabe einer Heirat, Scheidung oder den Tod des Partners vorkommt, wird als Ereignis der Person beachtet; ein solches Ereignis tritt aber, wie noch zu sehen sein wird, nicht ein.

Anfang und Ende einer Ehe

Die Anfangs- und Endpunkte der Ehen werden durch eine Jahresangabe angegeben und, falls dies möglich ist, durch die Angabe des Monats.

Die Erstellung dieser Daten ist recht kompliziert, wenn die Details betrachtet werden. So kann Anfang oder Ende einer retrospektiv angegebenen Ehe auch im 11-Jahres-Fenster liegen, falls die retrospektive Befragung später durchgeführt wurde. Fehler bei der Datenübertragung vom Fragebogen in die Datei oder irrige Angaben des Befragten, beides kann nicht immer auseinandergehalten werden, führen zu fehlenden Hochzeiten oder Bigamisten.

Zur Erstellung der Ehedaten aus der Datei *ehen.dat* habe ich ein C-Programm entwickelt. Innerhalb dieses Programms wird auf jeden möglichen Fehler reagiert, indem der gefundene Fehler als Fehlernummer und die Nummer der Person, bei der der Fehler aufgetreten ist, ausgegeben werden. Anhand dieser Liste kann nun der Fehler individuell behoben werden oder die Person wird als Problemfall gekennzeichnet. Die Markierung als Problemfall heißt dann: dieser Person konnten keine Ereignisdaten im Ehebereich zugeordnet werden, sie kann aber bei der Analyse anderer Daten, wie den Erwerbsspiels, durchaus herangezogen werden.

Anfänglich habe ich mich für die erste Methode entschieden und einzelne Personendaten angesehen, um Fehler individuell zu behandeln. Da jedoch sehr viele Personen Fehler aufweisen und sich die Bearbeitung dieser als sehr arbeitsaufwendig herausgestellt hat, ging ich bald zur zweiten Methode über.

Ehepartner

Die Generierung der Partnernummer zu einer gegebenen Ehe ist eine weitere Anforderung an das C-Programm, welches die Ehedaten erzeugt. Die Partnernummer kann nur zu den Ehen gegeben sein, die auch im 11-Jahres-Fenster vorkommen¹³.

Soll eine Partnernummer einer Ehe zugeordnet werden, so sind einige Bedingungen zu erfüllen:

- Die Partnernummer sollte während der Ehe, bzw. dem Teil der Ehe, der im 11-Jahres-Fenster verläuft, möglichst gleich bleiben.

¹³D.h. Ehen, die im 11-Jahres-Fenster geschlossen wurden, aber auch Ehen, die vor der monatsbezogenen Erhebung liegen und in das 11-Jahres-Fenster hineinreichen.

- Bezüglich der Partnernummer ist gefordert, daß der Partnerzeiger während der Ehedauer unverändert eine 1 oder 3 (eindeutiger Ehepartner bzw. vermutlicher Ehepartner) als Wert besitzt. Eine Ausnahme davon ist nur im ersten oder letzten Ehejahr erlaubt, da dann ein anderer Partner noch bzw. schon vorhanden sein kann.
- In diese Problematik hinein spielen auch die Trennungszeiten vor dem Eheende. Während dieser Zeiten kann ein anderer Partner vorliegen, obwohl die Ehe noch viele Jahre besteht.

Aus den geforderten Bedingungen und den möglichen Fehlern, ist sofort die Komplexität des C-Programms abzuleiten. Dieses sollte möglichst jeden Fehler erkennen und anzeigen. Auch hier gibt das C-Programm wie bereits bei den Ehe-daten eine Fehlerliste mit der Personennummer und der Partnernummer aus. Mit den Erfahrungen bzgl. des Aufwandes einer Kontrolle der individuellen Fehler, wurden hier alle Personen mit Fehlerausgaben als Problemfall markiert.

Die Problemfälle

Das C-Programm hat im ersten Durchlauf 294 Personen mit den unterschiedlichsten Fehlern ausgegeben. Von diesen Personen habe ich 93 Personen durch Zuweisungen bzw. Veränderungen von Daten auf individueller Ebene¹⁴ korrigiert:

- bei 6 Personen wurde eine Heirat hinzugefügt;
- bei 8 Personen wurde der Tod des Partners entfernt;
- bei 17 Personen wurde eine Scheidung hinzugefügt;
- bei 20 Personen wurde eine Heirat entfernt;
- bei 19 Personen wurde eine Scheidung entfernt;
- bei 1 Person wurde die Scheidung verschoben;
- bei 1 Person wurden alle Ereignisse entfernt;
- bei 14 Personen mußte das Ende der Ehe eingetragen werden. Da aber das Jahr nicht eindeutig identifiziert werden konnte, habe ich das Label "unbekanntes Ende einer Ehe" eingeführt und für die Personen verwandt;
- bei 9 Personen stellte sich dasselbe Problem bezgl. des Anfangs einer Ehe: ihnen mußte eine Heirat durch das Label "unbekannter Anfang einer Ehe" hinzugefügt werden.

¹⁴Diese Korrekturen hat das C-Programm durchgeführt; ich habe nicht die Daten direkt verändert.

Die übrigen Personen wurden, als Problemfälle gekennzeichnet, in die Datei *ehe_ereignisse.dat* übernommen.

Statistische Informationen zur Datei *ehe_ereignisse.dat*

Das C-Programm erzeugte schließlich 13631 Einträge für die Datei *ehe_ereignisse.dat*¹⁵:

- 11106 Anfangspunkte von Ehen (Spalte EREIGNIS, Wert 110), mit:
 - 9648 retrospektiv angegebene Heiraten (Spalte ART, Wert=1);
 - 1449 Heiraten im 11-Jahres-Fenster¹⁶ (Spalte ART, Wert=8);
 - 9 Ereignisse des Labels "unbekannte Heirat" (Spalte ART, Wert=21);
- 2307 Endpunkte von Ehen (Spalte EREIGNIS, Wert 120);
 - 896 retrospektiv angegebene Scheidungen (Spalte ART, Wert=2);
 - 813 retrospektiv angegebener Tod des Ehepartners (Spalte ART, Wert=3);
 - 26 retrospektiv angegebene Endpunkte einer Ehe, ohne Angabe des Grundes (Spalte ART, Wert=4);
 - 252 Scheidungen im 11-Jahres-Fenster (Spalte ART, Wert=9);
 - 306 Todesfälle eines Partners (Spalte ART, Wert=10);
 - 8 Ereignisse des Labels "unbekannte Scheidung" (Spalte ART, Wert=22);
 - 2 Ereignisse des Labels "unbekannter Tod des Partners" (Spalte ART, Wert=23);
 - 4 Ereignisse des Labels "unbekanntes Ende einer Ehe" (Spalte ART, Wert=24);
- 201 Problemfälle (Spalte EREIGNIS, Wert 199);

4.3 Kinder

Zur Generierung der Geburtsdaten von Kindern stehen einerseits die retrospektiv erhobenen Daten aus dem Biographiekomplex und andererseits die Angaben der Eltern im 11-Jahres-Fenster zu den Geburten ihrer Kinder zur Verfügung. Beide Datenquellen werden zur Erzeugung der Datei *kinder.dat*, in die die Geburtsdaten der Kinder einfließen, herangezogen.

¹⁵Siehe auch Anhang E.

¹⁶Hierbei ist zu beachten, daß retrospektiv angegebene Heiraten, auch wenn sie im 11-Jahres-Fenster liegen, als retrospektiv angegebene Heiraten, Scheidungen etc. gelten.

4.3.1 Retrospektiv erhobene Daten über Kinder

Innerhalb der Biographiedatenerhebung wurden die Geburtsdaten von maximal acht Kindern retrospektiv erhoben. Die Herauslösung dieser Daten aus der Datei *biolela.dat*, in der die Biographidaten gegeben sind, ist eine einfache Selektion.

Problematisch und unvorteilhaft gestalten sich die Daten aufgrund der Art ihrer Erhebung. Die Geburtsdaten der Kinder wurden nur von den Müttern erhoben mit der Frage: "Haben Sie Kinder geboren". Vergessen wurde dabei zum einen der Vater, für den es somit keine retrospektiv erhobenen Daten über Kinder gibt, und andererseits Pflege- oder Adoptivkinder. Daß nur maximal acht Kinder retrospektiv erhoben wurden ist ein nicht abzuschätzendes Problem, denn immerhin gibt es 43 achte Kinder und selbst mit der Selektion auf Personen der Stichproben A und B (KSample 1 und 2) verbleiben 35 achte Kinder.

Innerhalb späterer Analysen, die mit den hier erzeugten Datensätzen arbeiten, können die Kinder einer Mutter einfach dem Vater, falls dieser bekannt ist, ebenfalls zugeordnet werden. Dazu ist die Verbindung über die Partnernummer, die für Ehen eingetragen ist, heranzuziehen. Dies ist zwar nicht ganz korrekt, dürfte aber dem Umstand, daß Kinder in einer Ehe von Bedeutung sind und das (Erwerbs-)Verhalten der Eltern beeinflussen, gerecht werden. Verloren sind auf jeden Fall alleinerziehende Väter, die jedoch vor 1984 eine Ausnahme darstellten und Ehepartner, deren Ehe nicht bis in das 11-Jahres-Fenster hineinreicht, so daß die Ehefrau also nicht befragt werden konnte.

Durch die Selektion der Geburten aus der Datei *biolela.dat* ergibt sich ein Teil der Daten der Datei *kinder.dat*, zu der im folgenden Abschnitt einige statistische Daten angegeben werden¹⁷. Bereits mit der Übernahme der Geburten in die Datei *kinder.dat* erfolgte die Selektion auf Personen der Stichproben A und B.

Einige statistische Daten zur Datei *kinder.dat*

Durch die Selektion auf die Stichproben A und B verbleiben 16212 Personen in der Datei, 7975 Männer und 8237 Frauen. Einen Überblick über die retrospektiv angegebenen Kindergeburten gibt die Tabelle 4.4.

¹⁷Siehe dazu auch Abschnitt D.

Tabelle 4.4: Auszählung der retrospektiv erhobenen Daten über Kinder in

Kind	Anzahl	<i>kinder.dat</i>	
		frühestes Geburtsjahr	spätestes Geburtsjahr
1	4333	1913	1994
2	2977	1921	1993
3	1306	1923	1994
4	570	1930	1988
5	259	1932	1984
6	132	1933	1983
7	66	1938	1983
8	35	1945	1977

Die Reihenfolge der retrospektiv angegebenen Kindergeburten ist nicht bei allen Personen aufsteigend nach den Geburtsdaten sortiert. Es gibt 33 Personen, bei denen die Geburten der Kinder nicht sortiert sind. Dies ist nicht problematisch, muß jedoch bei der späteren Bearbeitung der Daten beachtet werden (Kinderanzahl bezogen auf die Jahre).

Interessanter und für die weitere Datenaufbereitung ebenso beachtenswert sind Mehrlingsgeburten. Für die retrospektive Erhebung der Kindergeburten bedeutet dies mehr als ein Kind in einem Jahr. Da nicht auseinandergehalten werden kann, ob es sich um eine tatsächliche Mehrlingsgeburt oder Geburten in verschiedenen Monaten handelt, werden diese Fälle als Mehrlingsgeburten behandelt. Es gibt 79 Personen, bei denen Mehrlingsgeburten gegeben sind.

4.3.2 Geburten im 11-Jahres-Fenster

Innerhalb des 11-Jahres-Fensters werden die Geburten monatsbezogen erhoben. Die Daten liegen in den Dateien *ap.dat*, *bp.dat* bis *kp.dat* in jeweils zwei Spalten pro Jahr vor¹⁸. Wie bereits bei den Ehedaten werde ich die erste Spalte als höhere Priorität¹⁹ zuerst verwenden. Zu beachten ist hierbei, daß bei der Datenerhebung nur eine Geburt pro Jahr erfaßt wurde, Mehrlingsgeburten also nicht vorkommen können.

Die Daten aus den Dateien *ap.dat*, *bp.dat* bis *kp.dat* werden in die Datei *kinder.dat* übernommen. Die Tabelle 4.5 gibt einen Überblick über die Anzahl der Geburten in den einzelnen Wellen²⁰.

¹⁸Es erfolgte auch ein Erhebung von Geburten auf der Haushaltsebene. Da jedoch diese Kinder keiner Person zugeordnet werden können, verbleibt als Quelle für Kindergeburten die Erhebung innerhalb des Personenfragebogens.

¹⁹Dies mit der Annahme, da sich Personen besser an kürzer zurückliegende Ereignisse erinnern können.

²⁰Siehe auch Anhang D.

Tabelle 4.5: Auszählung der Geburten des 11-Jahres-Fensters in *biolela.dat*

Jahr	Anzahl Geburten
1984	300
1985	333
1986	367
1987	325
1988	377
1989	325
1990	382
1991	320
1992	335
1993	326
1994	75

Die weitere Verarbeitung übernimmt ein C-Programm, welches ausgehend von den retrospektiv erhobenen Kinderdaten diese um die Daten der Geburten des 11-Jahres-Fensters erweitert. Kindergeburten, die zwar im Zuge der Biographiedatenerhebung erhoben wurden, können gleichzeitig innerhalb des 11-Jahres-Fensters liegen. Das Programm muß in diesem Fall die Monatsangaben den Geburten zuweisen und darf keine weitere Geburt eintragen.

Mit der Überarbeitung der Kindergeburten entsteht die Datei *kinder_ereignisse.dat*, in der nach der Struktur der Datei *ereignisse.dat*²¹ die Kindergeburten aufgelistet sind.

Einige statistische Daten zur Datei *kinder_ereignisse.dat*

Die Datei *kinder_ereignisse.dat* enthält 12756 Geburten²². Entsprechend der unglücklichen Erhebung, die auf Frauen ausgerichtet war, gibt es 11062 Rekords von Frauen und nur 1694 von Männern²³. Von 9292 Geburten ist der Monat nicht bekannt, die anderen verteilen sich recht gleichmäßig über das Jahr. Die maximale Kinderanzahl ist 10, jedoch muß dabei beachtet werden, daß die retrospektive Befragung maximal acht Kinder in Betracht gezogen hatte.

4.4 Zusammenführung der Ereignisdaten

Die beiden fertigen Dateien, *ehe_ereignisse.dat* und *kinder_ereignisse.dat*, besitzen dieselbe Struktur und können somit einfach aneinandergehängt werden. Jedes Betriebssystem eines Computers stellt dazu einen einfachen Befehl zur Verfügung. Durch diese Zusammenführung entsteht die Datei *ereignisse.dat*, die

²¹Siehe Abschnitt 4.1.

²²Siehe auch Anhang E.

²³Wenn man bedenkt, daß 3465 Geburten im 11-Jahres-Fenster angegeben sind, ist auch dieses merkwürdig. Von mir konnte aber keine Ursache ausgemacht werden.

noch sortiert werden kann, um die Daten einer Person an einer Stelle zu kummulieren und in der richtigen zeitlichen Reihenfolge abzulegen. Dadurch kann der Lebenslauf einer Person bezüglich familiärer Daten sehr einfach verfolgt werden.

Anhang A

Statistische Kennzahlen zur Datei *erwerb.dat*

In diesem Abschnitt werden einige Informationen zur Erstellung der Datei *erwerb.dat* gegeben.

Aus der Datei *artkalen.dat* entsteht mit dem TDA-commandfile *artkalen.cf* die Datei *artkalen_spells.dat*. Analog entsteht aus der Datei *pbiospe.dat* mittels des TDA-commandfiles *pbiospe.cf* die Datei *pbiospe_spells.dat*. Beide Dateien, *artkalen_spells.dat* und *pbiospe_spells.dat*, können mit einem einfachen Befehl aneinandergehängt werden, da beide Dateien dieselbe Struktur aufweisen.

A.1 *artkalen.cf*

```

section = "Einlesen aus der Datei artkalen.dat";
dfile   = artkalen.dat;
noc     = 87170;

HW [8.0] = c1;
PW [8.0] = c2;
SW [2.0] = c3;
ORG [2.0] = c4;
TS [4.0] = c5;
TF [4.0] = c6;
CEWS [2.0] = c7;
KS [1.0] = c8;

#-----
section="Matchen mit Daten aus ppfad.dat";
dfile = ppfad.dat;
noc   = 30000;
PW1 [8.0] = c1;
SEX [1.0] = c2;
GEB [4.0] = c3;
KS1 [1.0] = c4;
match=PW1,PW;
#-----
section="Bestimme neue Variable";

ORG_NEU [2.0] = if eq(ORG,1) then 1 else
                if eq(ORG,2) then 1 else
                if eq(ORG,3) then 2 else
                if eq(ORG,4) then 3 else
                if eq(ORG,5) then 4 else
                if eq(ORG,6) then 5 else
                if eq(ORG,7) then 9 else
                if eq(ORG,8) then 6 else
                if eq(ORG,9) then 7 else
                if eq(ORG,10) then 8 else
                if eq(ORG,11) then 9 else
                if eq(ORG,12) then 9 else 0;
DES [2.0] = 0;
TS_NEU [4.0] = if le(TS,0) then 0 else ((83*12)+TS);
TF_NEU [4.0] = if le(TF,0) then 0 else ((83*12)+TF);
DURATION [4.0] = TF_NEU-TS_NEU+1;
MARKER [1.0] = 1;

```

```

tsel = ge(GEB,1900).(eq(KS1,1)+eq(KS1,2)); Selektionen

keep = PN,SEX,GEB,KS1,SN,ORG_NEU,DES,TS_NEU,TF_NEU,CENS,
      DURATION,MARKER;
pdata = artkalen_spells.dat;
dtdata = artkalen_spells.cf;

```

A.2 *pbiospe.cf*

```

section="Einlesen aus der Datei pbiospe.dat";

dfile = pbiospe.dat;
noc = 96768;

PN [8.0] = c2;
SN = c3;
ORG [2.0] = c4;
TS [4.0] = c5;
TF [4.0] = c6;
CENS [3.0] = c11;
#-----
section="Matchen mit Daten aus ppfad.dat";
dfile = ppfad.dat;
noc = 30000;
PN1 [8.0] = c1;
SEX [1.0] = c2;
GEB [4.0] = c3;
KS [1.0] = c4;
match=PN1,PN;
#-----
section="Berechnen neuer Variablen";
tsel=(eq(KS,1)+eq(KS,2)).ge(GEB,1900); Selektionen

ORG_NEU [2.0] = if eq(ORG,1) then 6 else
                if eq(ORG,2) then 3 else
                if eq(ORG,3) then 7 else
                if eq(ORG,4) then 1 else
                if eq(ORG,5) then 2 else
                if eq(ORG,6) then 4 else
                if eq(ORG,7) then 8 else
                if eq(ORG,8) then 5 else
                if eq(ORG,9) then 9 else 0;
DES [2.0] = 0;
TS_NEU [2.0] = if le(TS,0) then 0 else
               if gt(TS,65) then -1 else ((GEB-1900+TS)*12+1);
TF_NEU [2.0] = if le(TF,0) then 0 else
               if gt(TF,65) then -1 else ((GEB-1900+TF)*12+12);
DURATION [2.0] = TF_NEU-TS_NEU+1;
MARKER [2.0] = 2;

keep = PN,SEX,GEB,KS,SN,ORG_NEU,DES,TS_NEU,TF_NEU,CENS,DURATION,MARKER;
pdata = pbiospe_spells.dat;

```

```
dtdd = pbiospe_spells.cf;
```

A.3 *artkalen_spells.dat*

Aus Gründen der Platzersparnis wurden einige Tabellen gekürzt. Die Stellen sind mit Punkten (...) gekennzeichnet.

```
TDA. Analysis of Transition Data (6.01). Thu Aug 29 13:57:57 1996
Command(s): cf=spells.cf
```

```
-----
Module: tda_gd [data generation]. Current memory: 134799 bytes.
```

```
New variables: 12
Set maximum number of cases to: 70000
Will generate a new data matrix with a maximum of 70000 cases.
```

```
Data file(s): artkalen_spells.dat
Assuming free format data file(s).
```

Idx	Variable	T	S	PFmt	Definition
1	PN	3	4	8.0	c1
2	SEX	3	4	1.0	c2
3	GEB	3	4	4.0	c3
4	KS	3	4	1.0	c4
5	SN	3	4	0.0	c5
6	ORG	3	4	2.0	c6
7	DES	3	8	2.0	c7
8	TS	3	4	4.0	c8
9	TF	3	4	4.0	c9
10	CENS	3	4	0.0	c10
11	DURATION	3	4	4.0	c11
12	MARKER	3	8	1.0	c12

```
Reading data file: artkalen_spells.dat
```

```
Read records: 52036
Created data for 12 variable(s) and 52036 cases.
Created a new data matrix.
Records containing (other types of) missing values: 0
```

```
End of module tda_gd (4055034).
```

```
-----
Module: tda_ds [descriptive statistics]. Current memory: 4055034 bytes.
```

```
Descriptive statistics: dstat
Temporary case selection (tsel): not active.
Number of cases: 52036
```

Variable	Minimum	Maximum	Mean	Std.Dev.
PN	101.0000	965902.0000	343093.0019	202399.6390
SEX	1.0000	2.0000	1.5406	0.4984

GEB	1900.0000	1979.0000	1952.0290	16.7194
KS	1.0000	2.0000	1.2700	0.4440
SN	1.0000	35.0000	3.4904	3.1048
ORG	0.0000	9.0000	3.9520	2.9197
DES	0.0000	0.0000	0.0000	0.0000
TS	997.0000	1128.0000	1043.7694	42.1398
TF	997.0000	1128.0000	1071.4669	43.5000
CENS	-2.0000	9.0000	2.5491	2.0740
DURATION	1.0000	132.0000	28.6975	35.4239
MARKER	1.0000	1.0000	1.0000	0.0000

Frequency distribution: freq1=SEX,GEB,KS,ORG,TS,TF,DURATION,MARKER
Max number of categories: 52036

Index	SEX	Frequency	Pct	Cumulated	Pct
1	1	23906.00	45.94	23906.00	45.94
2	2	28130.00	54.06	52036.00	100.00

Sum 52036.00 100.00

Index	GEB	Frequency	Pct	Cumulated	Pct
1	1900	53.00	0.10	53.00	0.10
2	1901	36.00	0.07	89.00	0.17
3	1902	71.00	0.14	160.00	0.31
4	1903	51.00	0.10	211.00	0.41
5	1904	69.00	0.13	280.00	0.54
6	1905	78.00	0.15	358.00	0.69
7	1906	86.00	0.17	444.00	0.85

...

70	1969	1677.00	3.22	46593.00	89.54
71	1970	1360.00	2.61	47953.00	92.15
72	1971	1057.00	2.03	49010.00	94.18
73	1972	897.00	1.72	49907.00	95.91
74	1973	778.00	1.50	50685.00	97.40
75	1974	546.00	1.05	51231.00	98.45
76	1975	396.00	0.76	51627.00	99.21
77	1976	250.00	0.48	51877.00	99.69
78	1977	158.00	0.30	52035.00	100.00
79	1979	1.00	0.00	52036.00	100.00

Sum 52036.00 100.00

Index	KS	Frequency	Pct	Cumulated	Pct
1	1	37986.00	73.00	37986.00	73.00
2	2	14050.00	27.00	52036.00	100.00

Sum 52036.00 100.00

Index	ORG	Frequency	Pct	Cumulated	Pct
1	0	1921.00	3.69	1921.00	3.69

2	1	15356.00	29.51	17277.00	33.20
3	2	5451.00	10.48	22728.00	43.68
4	3	3554.00	6.83	26282.00	50.51
5	4	5017.00	9.64	31299.00	60.15
6	5	3544.00	6.81	34843.00	66.96
7	6	4841.00	9.30	39684.00	76.26
8	7	596.00	1.15	40280.00	77.41
9	8	8086.00	15.54	48366.00	92.95
10	9	3670.00	7.05	52036.00	100.00

Sum 52036.00 100.00

Index	TS	Frequency	Pct	Cumulated	Pct

1	997	13194.00	25.36	13194.00	25.36
2	998	112.00	0.22	13306.00	25.57
3	999	163.00	0.31	13469.00	25.88
4	1000	203.00	0.39	13672.00	26.27
5	1001	161.00	0.31	13833.00	26.58
6	1002	147.00	0.28	13980.00	26.87
7	1003	252.00	0.48	14232.00	27.35
8	1004	300.00	0.58	14532.00	27.93
9	1005	296.00	0.57	14828.00	28.50
10	1006	226.00	0.43	15054.00	28.93
11	1007	165.00	0.32	15219.00	29.25
12	1008	153.00	0.29	15372.00	29.54
13	1009	2553.00	4.91	17925.00	34.45
14	1010	119.00	0.23	18044.00	34.68
15	1011	190.00	0.37	18234.00	35.04
16	1012	209.00	0.40	18443.00	35.44
17	1013	162.00	0.31	18605.00	35.75
18	1014	172.00	0.33	18777.00	36.08
19	1015	264.00	0.51	19041.00	36.59
20	1016	375.00	0.72	19416.00	37.31
21	1017	298.00	0.57	19714.00	37.89
22	1018	282.00	0.54	19996.00	38.43
23	1019	157.00	0.30	20153.00	38.73
24	1020	145.00	0.28	20298.00	39.01
25	1021	1809.00	3.48	22107.00	42.48
26	1022	121.00	0.23	22228.00	42.72
27	1023	163.00	0.31	22391.00	43.03
28	1024	210.00	0.40	22601.00	43.43
29	1025	150.00	0.29	22751.00	43.72
30	1026	132.00	0.25	22883.00	43.98
31	1027	262.00	0.50	23145.00	44.48
32	1028	332.00	0.64	23477.00	45.12
33	1029	309.00	0.59	23786.00	45.71
34	1030	247.00	0.47	24033.00	46.19
35	1031	158.00	0.30	24191.00	46.49
36	1032	122.00	0.23	24313.00	46.72
37	1033	1567.00	3.01	25880.00	49.73
38	1034	121.00	0.23	26001.00	49.97
39	1035	171.00	0.33	26172.00	50.30
40	1036	192.00	0.37	26364.00	50.66
41	1037	158.00	0.30	26522.00	50.97

42	1038	151.00	0.29	26673.00	51.26
...					
118	1114	259.00	0.50	48491.00	93.19
119	1115	140.00	0.27	48631.00	93.46
120	1116	115.00	0.22	48746.00	93.68
121	1117	1198.00	2.30	49944.00	95.98
122	1118	124.00	0.24	50068.00	96.22
123	1119	176.00	0.34	50244.00	96.56
124	1120	259.00	0.50	50503.00	97.05
125	1121	152.00	0.29	50655.00	97.35
126	1122	149.00	0.29	50804.00	97.63
127	1123	250.00	0.48	51054.00	98.11
128	1124	282.00	0.54	51336.00	98.65
129	1125	233.00	0.45	51569.00	99.10
130	1126	222.00	0.43	51791.00	99.53
131	1127	138.00	0.27	51929.00	99.79
132	1128	107.00	0.21	52036.00	100.00

Sum		52036.00	100.00		
-----	--	----------	--------	--	--

Index	TF	Frequency	Pct	Cumulated	Pct
1	997	102.00	0.20	102.00	0.20
2	998	151.00	0.29	253.00	0.49
3	999	197.00	0.38	450.00	0.86
4	1000	153.00	0.29	603.00	1.16
5	1001	141.00	0.27	744.00	1.43
6	1002	244.00	0.47	988.00	1.90
7	1003	285.00	0.55	1273.00	2.45
8	1004	273.00	0.52	1546.00	2.97
9	1005	224.00	0.43	1770.00	3.40
10	1006	153.00	0.29	1923.00	3.70
11	1007	157.00	0.30	2080.00	4.00
12	1008	3343.00	6.42	5423.00	10.42
13	1009	114.00	0.22	5537.00	10.64
14	1010	177.00	0.34	5714.00	10.98
15	1011	206.00	0.40	5920.00	11.38
16	1012	155.00	0.30	6075.00	11.67
17	1013	171.00	0.33	6246.00	12.00
18	1014	278.00	0.53	6524.00	12.54
19	1015	365.00	0.70	6889.00	13.24
20	1016	275.00	0.53	7164.00	13.77
21	1017	247.00	0.47	7411.00	14.24
22	1018	162.00	0.31	7573.00	14.55
23	1019	145.00	0.28	7718.00	14.83
24	1020	2497.00	4.80	10215.00	19.63
25	1021	122.00	0.23	10337.00	19.87
26	1022	156.00	0.30	10493.00	20.16
27	1023	216.00	0.42	10709.00	20.58
28	1024	163.00	0.31	10872.00	20.89
29	1025	131.00	0.25	11003.00	21.14
30	1026	267.00	0.51	11270.00	21.66
31	1027	328.00	0.63	11598.00	22.29
32	1028	292.00	0.56	11890.00	22.85
33	1029	227.00	0.44	12117.00	23.29

34	1030	173.00	0.33	12290.00	23.62
35	1031	121.00	0.23	12411.00	23.85
36	1032	1831.00	3.52	14242.00	27.37
37	1033	113.00	0.22	14355.00	27.59
38	1034	163.00	0.31	14518.00	27.90
39	1035	194.00	0.37	14712.00	28.27
40	1036	149.00	0.29	14861.00	28.56
41	1037	141.00	0.27	15002.00	28.83
42	1038	266.00	0.51	15268.00	29.34
43	1039	297.00	0.57	15565.00	29.91
...					
123	1119	294.00	0.56	40368.00	77.58
124	1120	161.00	0.31	40529.00	77.89
125	1121	150.00	0.29	40679.00	78.17
126	1122	267.00	0.51	40946.00	78.69
127	1123	292.00	0.56	41238.00	79.25
128	1124	217.00	0.42	41455.00	79.67
129	1125	222.00	0.43	41677.00	80.09
130	1126	142.00	0.27	41819.00	80.37
131	1127	136.00	0.26	41955.00	80.63
132	1128	10081.00	19.37	52036.00	100.00

Sum		52036.00	100.00		

Index	MARKER	Frequency	Pct	Cumulated	Pct

1	1	52036.00	100.00	52036.00	100.00

Sum		52036.00	100.00		

End of module tda_ds (4055034).

End of program. Memory used: 6553130 bytes. Thu Aug 29 13:58:38 1996

A.4 *pbiospe_spells.dat*

Aus Gründen der Platzersparnis wurden einige Tabellen gekürzt. Die Stellen sind mit Punkten (...) gekennzeichnet.

TDA. Analysis of Transition Data (6.01). Thu Aug 29 14:16:19 1996
Command(s): cf=spells.cf

Module: tda_gd [data generation]. Current memory: 134798 bytes.

New variables: 12

Set maximum number of cases to: 70000

Will generate a new data matrix with a maximum of 70000 cases.

Data file(s): pbiospe_spells.dat

Assuming free format data file(s).

Idx Variable T S PFmt Definition

```

1 PH      3 4   8.0 c1
2 SEX     3 4   1.0 c2
3 GEB     3 4   4.0 c3
4 KS      3 4   1.0 c4
5 SN      3 4   0.0 c5
6 ORG     3 4   2.0 c6
7 DES     3 8   2.0 c7
8 TS      3 4   4.0 c8
9 TF      3 4   4.0 c9
10 CENS    3 4   0.0 c10
11 DURATION 3 4   4.0 c11
12 MARKER  3 8   1.0 c12

```

Reading data file: /data/userdat/beate/daten/sfb/pbiospe_spells.dat

Read records: 70000

Created data for 12 variable(s) and 70000 cases.

Created a new data matrix.

Records containing (other types of) missing values: 0

End of module tda_gd (4055033).

Module: tda_ds [descriptive statistics]. Current memory: 4055033 bytes.

Descriptive statistics: dstat

Temporary case selection (tsel): not active.

Number of cases: 70000

Variable	Minimum	Maximum	Mean	Std.Dev.
PH	101.0000	965902.0000	321448.0362	202955.9558
SEX	1.0000	2.0000	1.5325	0.4989
GEB	1900.0000	1979.0000	1945.9306	17.8755
KS	1.0000	2.0000	1.2070	0.4052
SN	1.0000	19.0000	3.3838	2.1924
ORG	1.0000	9.0000	4.2052	2.7140
DES	0.0000	0.0000	0.0000	0.0000
TS	181.0000	1501.0000	853.2782	214.9296
TF	192.0000	1572.0000	931.9577	201.6138
CENS	1.0000	4.0000	1.8111	0.9238
DURATION	12.0000	612.0000	79.6795	109.2729
MARKER	2.0000	2.0000	2.0000	0.0000

Frequency distribution: freq1=SEX,GEB,KS,ORG,TS,TF,DURATION,MARKER
Max number of categories: 70000

Index	SEX	Frequency	Pct	Cumulated	Pct
1	1	32728.00	46.75	32728.00	46.75
2	2	37272.00	53.25	70000.00	100.00

Sum		70000.00	100.00		

Index	GEB	Frequency	Pct	Cumulated	Pct
-------	-----	-----------	-----	-----------	-----

1	1900	124.00	0.18	124.00	0.18
2	1901	116.00	0.17	240.00	0.34
3	1902	181.00	0.26	421.00	0.60
4	1903	124.00	0.18	545.00	0.78
5	1904	180.00	0.26	725.00	1.04
6	1905	190.00	0.27	915.00	1.31
7	1906	229.00	0.33	1144.00	1.63
8	1907	294.00	0.42	1438.00	2.05
9	1908	294.00	0.42	1732.00	2.47
10	1909	371.00	0.53	2103.00	3.00
...					
73	1972	620.00	0.89	68460.00	97.80
74	1973	514.00	0.73	68974.00	98.53
75	1974	385.00	0.55	69359.00	99.08
76	1975	297.00	0.42	69656.00	99.51
77	1976	191.00	0.27	69847.00	99.78
78	1977	152.00	0.22	69999.00	100.00
79	1979	1.00	0.00	70000.00	100.00

Sum 70000.00 100.00

Index	KS	Frequency	Pct	Cumulated	Pct
1	1	55507.00	79.30	55507.00	79.30
2	2	14493.00	20.70	70000.00	100.00

Sum 70000.00 100.00

Index	ORG	Frequency	Pct	Cumulated	Pct
1	1	18025.00	25.75	18025.00	25.75
2	2	5748.00	8.21	23773.00	33.96
3	3	10864.00	15.52	34637.00	49.48
4	4	5352.00	7.65	39989.00	57.13
5	5	2747.00	3.92	42736.00	61.05
6	6	10861.00	15.52	53597.00	76.57
7	7	3693.00	5.28	57290.00	81.84
8	8	8300.00	11.86	65590.00	93.70
9	9	4410.00	6.30	70000.00	100.00

Sum 70000.00 100.00

Index	TS	Frequency	Pct	Cumulated	Pct
1	181	36.00	0.05	36.00	0.05
2	193	38.00	0.05	74.00	0.11
3	205	50.00	0.07	124.00	0.18
4	217	56.00	0.08	180.00	0.26
5	229	67.00	0.10	247.00	0.35
6	241	71.00	0.10	318.00	0.45
7	253	98.00	0.14	416.00	0.59
8	265	112.00	0.16	528.00	0.75
9	277	127.00	0.18	655.00	0.94
10	289	153.00	0.22	808.00	1.15

11	301	165.00	0.24	973.00	1.39
12	313	165.00	0.24	1138.00	1.63
13	325	195.00	0.28	1333.00	1.90
14	337	210.00	0.30	1543.00	2.20
...					
73	1045	2334.00	3.33	57542.00	82.20
74	1057	2320.00	3.31	59862.00	85.52
75	1069	2215.00	3.16	62077.00	88.68
76	1081	2078.00	2.97	64155.00	91.65
77	1093	2002.00	2.86	66157.00	94.51
78	1105	2135.00	3.05	68292.00	97.56
79	1117	1689.00	2.41	69981.00	99.97
80	1129	13.00	0.02	69994.00	99.99
81	1141	2.00	0.00	69996.00	99.99
82	1153	1.00	0.00	69997.00	100.00
83	1189	1.00	0.00	69998.00	100.00
84	1441	1.00	0.00	69999.00	100.00
85	1501	1.00	0.00	70000.00	100.00

Sum 70000.00 100.00

Index	TF	Frequency	Pct	Cumulated	Pct
1	192	7.00	0.01	7.00	0.01
2	204	10.00	0.01	17.00	0.02
3	216	18.00	0.03	35.00	0.05
4	228	23.00	0.03	58.00	0.08
5	240	24.00	0.03	82.00	0.12
6	252	44.00	0.06	126.00	0.18
7	264	44.00	0.06	170.00	0.24
8	276	60.00	0.09	230.00	0.33
9	288	59.00	0.08	289.00	0.41
10	300	72.00	0.10	361.00	0.52
11	312	88.00	0.13	449.00	0.64
12	324	96.00	0.14	545.00	0.78
...					
83	1188	2.00	0.00	69993.00	99.99
84	1200	2.00	0.00	69995.00	99.99
85	1224	1.00	0.00	69996.00	99.99
86	1236	1.00	0.00	69997.00	100.00
87	1272	1.00	0.00	69998.00	100.00
88	1452	1.00	0.00	69999.00	100.00
89	1572	1.00	0.00	70000.00	100.00

Sum 70000.00 100.00

Index	MARKER	Frequency	Pct	Cumulated	Pct
1	2	70000.00	100.00	70000.00	100.00

Sum 70000.00 100.00

End of module tda_ds (4055033).

End of program. Memory used: 7415417 bytes. Thu Aug 29 14:17:12 1996

Anhang B

Statistische Kennzahlen zur Datei *erwerb_spells.dat*

In diesem Abschnitt werden statistische Informationen zur Datei *erwerbsspells.dat* bereitgestellt. Aus Gründen der Platzersparnis wurden einige Tabellen gekürzt. Diese Stellen sind mit Punkten (...) gekennzeichnet.

TDA. Analysis of Transition Data (6.01). Thu Aug 29 15:49:43 1996
Command(s): cf=erwerb_spells.cf

Module: tda_gd [data generation]. Current memory: 134775 bytes.

New variables: 13
Set maximum number of cases to: 200000
Will generate a new data matrix with a maximum of 200000 cases.

Data file(s): x.dat
Assuming free format data file(s).

Idx	Variable	Label	T	S	PFmt	Definition
1	V1	PH	3	4	0.0	c1
2	V2	SEX	3	4	0.0	c2
3	V3	GEB	3	4	0.0	c3
4	V4	KS	3	4	0.0	c4
5	V5	SW	3	4	0.0	c5
6	V6	ORG	3	4	0.0	c6
7	V7	DES	3	4	0.0	c7
8	V8	TS	3	4	0.0	c8
9	V9	TF	3	4	0.0	c9
10	V10	CEWS	3	4	0.0	c10
11	V11	DAUER	3	4	0.0	c11
12	V12	MARKER	3	4	0.0	c12
13	V13	LUECKEN	3	4	0.0	c13

Reading data file: x.dat

Read records: 63364
Created data for 13 variable(s) and 63364 cases.
Created a new data matrix.
Records containing (other types of) missing values: 0

End of module tda_gd (10535074).

Module: tda_ds [descriptive statistics]. Current memory: 10535074 bytes.

Descriptive statistics: dstat
Temporary case selection (tsel): not active.
Number of cases: 63364

Variable	Label	Minimum	Maximum	Mean	Std.Dev.
V1	PH	101.0000	965902.0000	332592.3686	206481.3742
V2	SEX	1.0000	2.0000	1.5367	0.4987
V3	GEB	1900.0000	1977.0000	1944.7506	18.5722
V4	KS	1.0000	2.0000	1.2445	0.4298
V5	SW	1.0000	25.0000	2.9870	1.9542
V6	ORG	1.0000	9.0000	3.8662	2.6071

V7	DES	1.0000	9.0000	3.4067	2.6558
V8	TS	181.0000	1129.0000	836.7033	214.3373
V9	TF	192.0000	1140.0000	920.1229	205.1277
V10	CENS	0.0000	3.0000	0.7839	0.8933
V11	DAUER	1.0000	768.0000	84.4196	112.1690
V12	MARKER	1.0000	3.0000	1.9854	0.3986
V13	LUECKEN	0.0000	12.0000	0.2209	1.3205

Frequency distribution: freq1=V2,V3,V4,V5,V6,V7,V8,V9,V10,V11,V12,V13
Max number of categories: 63364

Index	V2	Frequency	Pct	Cumulated	Pct
1	1	29359.00	46.33	29359.00	46.33
2	2	34005.00	53.67	63364.00	100.00

Sum 63364.00 100.00

Index	V3	Frequency	Pct	Cumulated	Pct
1	1900	155.00	0.24	155.00	0.24
2	1901	136.00	0.21	291.00	0.46
3	1902	225.00	0.36	516.00	0.81
4	1903	158.00	0.25	674.00	1.06
5	1904	230.00	0.36	904.00	1.43
6	1905	235.00	0.37	1139.00	1.80
7	1906	282.00	0.45	1421.00	2.24
8	1907	372.00	0.59	1793.00	2.83
9	1908	349.00	0.55	2142.00	3.38
10	1909	426.00	0.67	2568.00	4.05
11	1910	475.00	0.75	3043.00	4.80
12	1911	466.00	0.74	3509.00	5.54
13	1912	491.00	0.77	4000.00	6.31
14	1913	489.00	0.77	4489.00	7.08

...

69	1968	1278.00	2.02	58473.00	92.28
70	1969	1115.00	1.76	59588.00	94.04
71	1970	909.00	1.43	60497.00	95.48
72	1971	765.00	1.21	61262.00	96.68
73	1972	586.00	0.92	61848.00	97.61
74	1973	505.00	0.80	62353.00	98.40
75	1974	362.00	0.57	62715.00	98.98
76	1975	291.00	0.46	63006.00	99.44
77	1976	204.00	0.32	63210.00	99.76
78	1977	154.00	0.24	63364.00	100.00

Sum 63364.00 100.00

Index	V4	Frequency	Pct	Cumulated	Pct
1	1	47873.00	75.55	47873.00	75.55
2	2	15491.00	24.45	63364.00	100.00

Sum 63364.00 100.00

Index	V5	Frequency	Pct	Cumulated	Pct
1	1	16119.00	25.44	16119.00	25.44
2	2	14857.00	23.45	30976.00	48.89
3	3	12129.00	19.14	43105.00	68.03
4	4	8446.00	13.33	51551.00	81.36
5	5	5278.00	8.33	56829.00	89.69
6	6	3038.00	4.79	59867.00	94.48
7	7	1664.00	2.63	61531.00	97.11
8	8	856.00	1.35	62387.00	98.46
9	9	447.00	0.71	62834.00	99.16
10	10	229.00	0.36	63063.00	99.52
11	11	121.00	0.19	63184.00	99.72
12	12	65.00	0.10	63249.00	99.82
13	13	40.00	0.06	63289.00	99.88
14	14	26.00	0.04	63315.00	99.92
15	15	15.00	0.02	63330.00	99.95
16	16	10.00	0.02	63340.00	99.96
17	17	7.00	0.01	63347.00	99.97
18	18	6.00	0.01	63353.00	99.98
19	19	3.00	0.00	63356.00	99.99
20	20	2.00	0.00	63358.00	99.99
21	21	2.00	0.00	63360.00	99.99
22	22	1.00	0.00	63361.00	100.00
23	23	1.00	0.00	63362.00	100.00
24	24	1.00	0.00	63363.00	100.00
25	25	1.00	0.00	63364.00	100.00

Sum 63364.00 100.00

Index	V6	Frequency	Pct	Cumulated	Pct
1	1	18944.00	29.90	18944.00	29.90
2	2	5712.00	9.01	24656.00	38.91
3	3	10103.00	15.94	34759.00	54.86
4	4	3105.00	4.90	37864.00	59.76
5	5	3317.00	5.23	41181.00	64.99
6	6	10533.00	16.62	51714.00	81.61
7	7	3212.00	5.07	54926.00	86.68
8	8	6326.00	9.98	61252.00	96.67
9	9	2112.00	3.33	63364.00	100.00

Sum 63364.00 100.00

Index	V7	Frequency	Pct	Cumulated	Pct
1	1	25237.00	39.83	25237.00	39.83
2	2	6858.00	10.82	32095.00	50.65
3	3	7156.00	11.29	39251.00	61.95
4	4	3472.00	5.48	42723.00	67.42
5	5	5793.00	9.14	48516.00	76.57
6	6	2261.00	3.57	50777.00	80.14
7	7	3172.00	5.01	53949.00	85.14
8	8	7919.00	12.50	61868.00	97.64

9	9	1496.00	2.36	63364.00	100.00
---	---	---------	------	----------	--------

Sum	63364.00 100.00

Index	V8	Frequency	Pct	Cumulated	Pct

1	181	34.00	0.05	34.00	0.05
2	193	37.00	0.06	71.00	0.11
3	205	47.00	0.07	118.00	0.19
4	217	54.00	0.09	172.00	0.27
5	229	65.00	0.10	237.00	0.37
6	241	70.00	0.11	307.00	0.48
7	253	94.00	0.15	401.00	0.63
8	265	103.00	0.16	504.00	0.80
9	277	123.00	0.19	627.00	0.99
10	289	147.00	0.23	774.00	1.22
11	301	157.00	0.25	931.00	1.47
12	307	2.00	0.00	933.00	1.47
13	313	165.00	0.26	1098.00	1.73
14	325	188.00	0.30	1286.00	2.03
15	337	189.00	0.30	1475.00	2.33
16	349	196.00	0.31	1671.00	2.64
17	361	218.00	0.34	1889.00	2.98
18	373	190.00	0.30	2079.00	3.28
19	385	189.00	0.30	2268.00	3.58
20	397	220.00	0.35	2488.00	3.93
21	409	235.00	0.37	2723.00	4.30
22	421	287.00	0.45	3010.00	4.75
23	433	309.00	0.49	3319.00	5.24
24	445	344.00	0.54	3663.00	5.78
25	457	339.00	0.54	4002.00	6.32
26	469	431.00	0.68	4433.00	7.00
27	481	513.00	0.81	4946.00	7.81
28	487	1.00	0.00	4947.00	7.81
29	493	469.00	0.74	5416.00	8.55
30	499	1.00	0.00	5417.00	8.55
31	505	461.00	0.73	5878.00	9.28
32	517	523.00	0.83	6401.00	10.10
33	529	509.00	0.80	6910.00	10.91
34	541	696.00	1.10	7606.00	12.00
35	553	895.00	1.41	8501.00	13.42
36	559	2.00	0.00	8503.00	13.42
37	565	705.00	1.11	9208.00	14.53
38	577	687.00	1.08	9895.00	15.62
39	583	1.00	0.00	9896.00	15.62
40	589	643.00	1.01	10539.00	16.63
41	595	1.00	0.00	10540.00	16.63
42	601	676.00	1.07	11216.00	17.70
43	613	630.00	0.99	11846.00	18.70
44	625	658.00	1.04	12504.00	19.73
45	637	759.00	1.20	13263.00	20.93
46	643	1.00	0.00	13264.00	20.93
47	649	721.00	1.14	13985.00	22.07
48	661	797.00	1.26	14782.00	23.33
49	673	753.00	1.19	15535.00	24.52

50	679	1.00	0.00	15536.00	24.52
51	685	795.00	1.25	16331.00	25.77
52	697	742.00	1.17	17073.00	26.94
53	709	819.00	1.29	17892.00	28.24
...					
194	1098	18.00	0.03	60358.00	95.26
195	1099	54.00	0.09	60412.00	95.34
196	1100	51.00	0.08	60463.00	95.42
197	1101	37.00	0.06	60500.00	95.48
198	1102	21.00	0.03	60521.00	95.51
199	1103	7.00	0.01	60528.00	95.52
200	1104	7.00	0.01	60535.00	95.54
201	1105	1264.00	1.99	61799.00	97.53
202	1106	22.00	0.03	61821.00	97.56
203	1107	32.00	0.05	61853.00	97.62
204	1108	42.00	0.07	61895.00	97.68
205	1109	20.00	0.03	61915.00	97.71
206	1110	24.00	0.04	61939.00	97.75
207	1111	47.00	0.07	61986.00	97.83
208	1112	70.00	0.11	62056.00	97.94
209	1113	30.00	0.05	62086.00	97.98
210	1114	17.00	0.03	62103.00	98.01
211	1115	7.00	0.01	62110.00	98.02
212	1116	10.00	0.02	62120.00	98.04
213	1117	1010.00	1.59	63130.00	99.63
214	1118	18.00	0.03	63148.00	99.66
215	1119	22.00	0.03	63170.00	99.69
216	1120	34.00	0.05	63204.00	99.75
217	1121	18.00	0.03	63222.00	99.78
218	1122	18.00	0.03	63240.00	99.80
219	1123	31.00	0.05	63271.00	99.85
220	1124	33.00	0.05	63304.00	99.91
221	1125	17.00	0.03	63321.00	99.93
222	1126	8.00	0.01	63329.00	99.94
223	1127	4.00	0.01	63333.00	99.95
224	1128	3.00	0.00	63336.00	99.96
225	1129	28.00	0.04	63364.00	100.00

Sum		63364.00	100.00		

Index	V9	Frequency	Pct	Cumulated	Pct
1	192	8.00	0.01	8.00	0.01
2	204	6.00	0.01	14.00	0.02
3	216	21.00	0.03	35.00	0.06
4	228	20.00	0.03	55.00	0.09
5	240	25.00	0.04	80.00	0.13
6	252	41.00	0.06	121.00	0.19
7	264	39.00	0.06	160.00	0.25
8	276	52.00	0.08	212.00	0.33
9	288	60.00	0.09	272.00	0.43
10	300	68.00	0.11	340.00	0.54
11	306	2.00	0.00	342.00	0.54
12	312	85.00	0.13	427.00	0.67
13	324	97.00	0.15	524.00	0.83

14	336	94.00	0.15	618.00	0.98
15	348	103.00	0.16	721.00	1.14
16	360	124.00	0.20	845.00	1.33
17	372	124.00	0.20	969.00	1.53
18	384	132.00	0.21	1101.00	1.74
19	396	154.00	0.24	1255.00	1.98
20	408	161.00	0.25	1416.00	2.23
21	420	172.00	0.27	1588.00	2.51
22	432	162.00	0.26	1750.00	2.76
23	444	203.00	0.32	1953.00	3.08
24	456	203.00	0.32	2156.00	3.40
25	468	284.00	0.45	2440.00	3.85
26	480	346.00	0.55	2786.00	4.40
27	486	1.00	0.00	2787.00	4.40
28	492	306.00	0.48	3093.00	4.88
29	498	1.00	0.00	3094.00	4.88
30	504	285.00	0.45	3379.00	5.33
31	516	344.00	0.54	3723.00	5.88
...					
192	1096	25.00	0.04	50478.00	79.66
193	1097	18.00	0.03	50496.00	79.69
194	1098	54.00	0.09	50550.00	79.78
195	1099	51.00	0.08	50601.00	79.86
196	1100	37.00	0.06	50638.00	79.92
197	1101	21.00	0.03	50659.00	79.95
198	1102	7.00	0.01	50666.00	79.96
199	1103	7.00	0.01	50673.00	79.97
200	1104	1466.00	2.31	52139.00	82.28
201	1105	22.00	0.03	52161.00	82.32
202	1106	32.00	0.05	52193.00	82.37
203	1107	42.00	0.07	52235.00	82.44
204	1108	20.00	0.03	52255.00	82.47
205	1109	24.00	0.04	52279.00	82.51
206	1110	47.00	0.07	52326.00	82.58
207	1111	70.00	0.11	52396.00	82.69
208	1112	30.00	0.05	52426.00	82.74
209	1113	17.00	0.03	52443.00	82.76
210	1114	7.00	0.01	52450.00	82.78
211	1115	10.00	0.02	52460.00	82.79
212	1116	1661.00	2.62	54121.00	85.41
213	1117	18.00	0.03	54139.00	85.44
214	1118	22.00	0.03	54161.00	85.48
215	1119	34.00	0.05	54195.00	85.53
216	1120	18.00	0.03	54213.00	85.56
217	1121	18.00	0.03	54231.00	85.59
218	1122	31.00	0.05	54262.00	85.64
219	1123	33.00	0.05	54295.00	85.69
220	1124	17.00	0.03	54312.00	85.71
221	1125	8.00	0.01	54320.00	85.73
222	1126	4.00	0.01	54324.00	85.73
223	1127	3.00	0.00	54327.00	85.74
224	1128	8848.00	13.96	63175.00	99.70
225	1140	189.00	0.30	63364.00	100.00

Sum		63364.00	100.00		

Index	V10	Frequency	Pct	Cumulated	Pct
1	0	31930.00	50.39	31930.00	50.39
2	1	14877.00	23.48	46807.00	73.87
3	2	14877.00	23.48	61684.00	97.35
4	3	1680.00	2.65	63364.00	100.00

Sum 63364.00 100.00

Index	V12	Frequency	Pct	Cumulated	Pct
1	1	5503.00	8.68	5503.00	8.68
2	2	53282.00	84.09	58785.00	92.77
3	3	4579.00	7.23	63364.00	100.00

Sum 63364.00 100.00

Index	V13	Frequency	Pct	Cumulated	Pct
1	0	60511.00	95.50	60511.00	95.50
2	1	1054.00	1.66	61565.00	97.16
3	2	858.00	1.35	62423.00	98.51
4	3	10.00	0.02	62433.00	98.53
5	4	851.00	1.34	63284.00	99.87
6	5	17.00	0.03	63301.00	99.90
7	6	38.00	0.06	63339.00	99.96
8	10	2.00	0.00	63341.00	99.96
9	20	6.00	0.01	63347.00	99.97
10	40	6.00	0.01	63353.00	99.98
11	41	1.00	0.00	63354.00	99.98
12	100	4.00	0.01	63358.00	99.99
13	120	1.00	0.00	63359.00	99.99
14	200	5.00	0.01	63364.00	100.00

Sum 63364.00 100.00

End of module tda_ds (10535086).

End of program. Memory used: 14590838 bytes. Wed Sep 11 11:50:26 1996

Anhang C

Statistische Kennzahlen zur Datei *ehen.dat*

In diesem Abschnitt werden Informationen zur Erstellung der Datei *ehen.dat* bereitgestellt. Das TDA-commandfile *ehen.cf* setzt Daten aus den Dateien *biolela.dat*, *ppfad.dat* und *familie.dat* in die Datei *ehen.dat* um.

C.1 *ehen.cf*

```

section = "einlesen der Daten aus biolela.dat";
dfile=biolela.dat;
noc = 21747;
PW      [8.0] = c2;
ENE1A <1>[2.0] = c46;
ENE2A <1>[2.0] = c47;
ENE3A <1>[2.0] = c48;
ENE1S <1>[2.0] = c49;
ENE2S <1>[2.0] = c50;
ENE3S <1>[2.0] = c51;
ENE1E <1>[2.0] = c52;
ENE2E <1>[2.0] = c53;
ENE3E <1>[2.0] = c54;
#-----
section = "einlesen der Datei ppfad.dat";
dfile=ppfad.dat;
noc=27620;
PW1     [7.0] = c1;
SEX <1>[2.0] = c2;
GEB <2>[4.0] = c3;
KS  <1>[1.0] = c4;
match=PW1,PW; -> merge biolela + ppfad
#-----
section = "einlesen der Datei familie.dat";
dfile = familie.dat;
noc = 27620;
PW2 [7.0] = c1;
PW84 [6.0] = c5;
PW85 [6.0] = c6;
PW86 [6.0] = c7;
PW87 [6.0] = c8;
PW88 [6.0] = c9;
PW89 [6.0] = c10;
PW90 [7.0] = c11;
PW91 [7.0] = c12;
PW92 [7.0] = c13;
PW93 [7.0] = c14;
PW94 [7.0] = c15;
PZ84 <1>[2.0] = c16;
PZ85 <1>[2.0] = c17;
PZ86 <1>[2.0] = c18;
PZ87 <1>[2.0] = c19;
PZ88 <1>[2.0] = c20;
PZ89 <1>[2.0] = c21;
PZ90 <1>[2.0] = c22;
PZ91 <1>[2.0] = c23;
PZ92 <1>[2.0] = c24;

```



```

PZ93 <1>[2.0] = c25;
PZ94 <1>[2.0] = c26;
FAMST84 <1>[2.0] = c27;
FAMST85 <1>[2.0] = c28;
FAMST86 <1>[2.0] = c29;
FAMST87 <1>[2.0] = c30;
FAMST88 <1>[2.0] = c31;
FAMST89 <1>[2.0] = c32;
FAMST90 <1>[2.0] = c33;
FAMST91 <1>[2.0] = c34;
FAMST92 <1>[2.0] = c35;
FAMST93 <1>[2.0] = c36;
FAMST94 <1>[2.0] = c37;
HR84 [0.0] = c38;
HR85 [0.0] = c39;
HR86 [0.0] = c40;
HR87 [0.0] = c41;
HR88 [0.0] = c42;
HR89 [0.0] = c43;
HR90 [0.0] = c44;
HR91 [0.0] = c45;
HR92 [0.0] = c46;
HR93 [0.0] = c47;
HR94 [0.0] = c48;
SCH84 [0.0] = c49;
SCH85 [0.0] = c50;
SCH86 [0.0] = c51;
SCH87 [0.0] = c52;
SCH88 [0.0] = c53;
SCH89 [0.0] = c54;
SCH90 [0.0] = c55;
SCH91 [0.0] = c56;
SCH92 [0.0] = c57;
SCH93 [0.0] = c58;
SCH94 [0.0] = c59;
TOD84 [0.0] = c60;
TOD85 [0.0] = c61;
TOD86 [0.0] = c62;
TOD87 [0.0] = c63;
TOD88 [0.0] = c64;
TOD89 [0.0] = c65;
TOD90 [0.0] = c66;
TOD91 [0.0] = c67;
TOD92 [0.0] = c68;
TOD93 [0.0] = c69;
TOD94 [0.0] = c70;
match=PF2,PF;
#-----
section = "Ausgabe";
tsel=(eq(KS,1)+eq(KS,2)).ge(GEB,1900); Selektionen
pdata = ehcn.dat;

```

C.2 *ehen.dat*

Einige Tabellen mußten aus Platzgründen gekürzt werden. Diese Stellen sind mit Punkten (...) gekennzeichnet.

TDA. Analysis of Transition Data (6.01). Tue Sep 10 16:43:22 1996
Command(s): cf=ehen1.cf

Module: tda_gd [data generation]. Current memory: 135873 bytes.

New variables: 79

Set maximum number of cases to: 200000

Will generate a new data matrix with a maximum of 200000 cases.

Data file(s): ehen.dat

Assuming free format data file(s).

Idx	Variable	T	S	PFmt	Definition
1	PH	3	4	0.0	c1
2	SEX	3	4	0.0	c12
3	GEB	3	4	0.0	c13
4	KS	3	4	0.0	c14
38	HEIRAT84	3	4	0.0	c49
39	HEIRAT85	3	4	0.0	c50
40	HEIRAT86	3	4	0.0	c51
41	HEIRAT87	3	4	0.0	c52
42	HEIRAT88	3	4	0.0	c53
43	HEIRAT89	3	4	0.0	c54
44	HEIRAT90	3	4	0.0	c55
45	HEIRAT91	3	4	0.0	c56
46	HEIRAT92	3	4	0.0	c57
47	HEIRAT93	3	4	0.0	c58
48	HEIRAT94	3	4	0.0	c59
49	SCHEID84	3	4	0.0	c60
50	SCHEID85	3	4	0.0	c61
51	SCHEID86	3	4	0.0	c62
52	SCHEID87	3	4	0.0	c63
53	SCHEID88	3	4	0.0	c64
54	SCHEID89	3	4	0.0	c65
55	SCHEID90	3	4	0.0	c66
56	SCHEID91	3	4	0.0	c67
57	SCHEID92	3	4	0.0	c68
58	SCHEID93	3	4	0.0	c69
59	SCHEID94	3	4	0.0	c70
60	TOD84	3	4	0.0	c71
61	TOD85	3	4	0.0	c72
62	TOD86	3	4	0.0	c73
63	TOD87	3	4	0.0	c74
64	TOD88	3	4	0.0	c75
65	TOD89	3	4	0.0	c76
66	TOD90	3	4	0.0	c77
67	TOD91	3	4	0.0	c78
68	TOD92	3	4	0.0	c79
69	TOD93	3	4	0.0	c80

```

70 TOD94      3 4  0.0 c81
71 ANFANG1EHE 3 4  0.0 c2
72 ANFANG2EHE 3 4  0.0 c3
73 ANFANG3EHE 3 4  0.0 c4
74 STATUS1EHE 3 4  0.0 c5
75 STATUS2EHE 3 4  0.0 c6
76 STATUS3EHE 3 4  0.0 c7
77 ENDE1EHE   3 4  0.0 c8
78 ENDE2EHE   3 4  0.0 c9
79 ENDE3EHE   3 4  0.0 c10

```

Reading data file: shen.dat

Read records: 16131

Created data for 79 variable(s) and 16131 cases.

Created a new data matrix.

Records containing (other types of) missing values: 0

End of module tda_gd (63337697).

Module: tda_ds [descriptive statistics]. Current memory: 63337697 bytes.

Descriptive statistics: dstat

Temporary case selection (tsel): not active.

Number of cases: 16131

Variable	Minimum	Maximum	Mean	Std.Dev.
PN	101.0000	965902.0000	348172.7898	208398.3258
SEX	1.0000	2.0000	1.5073	0.5000
GEB	1900.0000	1979.0000	1947.7511	18.4066
KS	1.0000	2.0000	1.2807	0.4494
HEIRAT84	0.0000	12.0000	0.0858	0.8662
HEIRAT85	0.0000	12.0000	0.0958	0.8813
HEIRAT86	0.0000	12.0000	0.1152	0.9608
HEIRAT87	0.0000	12.0000	0.0946	0.8893
HEIRAT88	0.0000	12.0000	0.0867	0.8281
HEIRAT89	0.0000	12.0000	0.0967	0.8860
HEIRAT90	0.0000	12.0000	0.0857	0.8305
HEIRAT91	0.0000	12.0000	0.1138	0.9676
HEIRAT92	0.0000	12.0000	0.0932	0.8793
HEIRAT93	0.0000	12.0000	0.1004	0.9018
HEIRAT94	0.0000	9.0000	0.0099	0.2512
SCHEID84	0.0000	12.0000	0.0146	0.3508
SCHEID85	0.0000	12.0000	0.0127	0.3383
SCHEID86	0.0000	12.0000	0.0169	0.3774
SCHEID87	0.0000	11.0000	0.0092	0.2369
SCHEID88	0.0000	12.0000	0.0161	0.3744
SCHEID89	0.0000	12.0000	0.0138	0.3507
SCHEID90	0.0000	12.0000	0.0125	0.3235
SCHEID91	0.0000	12.0000	0.0183	0.4016
SCHEID92	0.0000	11.0000	0.0093	0.2533
SCHEID93	0.0000	12.0000	0.0177	0.3758
SCHEID94	0.0000	10.0000	0.0032	0.1252
TOD84	0.0000	12.0000	0.0158	0.3801

TOD85	0.0000	12.0000	0.0143	0.3487
TOD86	0.0000	12.0000	0.0187	0.3924
TOD87	0.0000	12.0000	0.0170	0.3764
TOD88	0.0000	12.0000	0.0172	0.3942
TOD89	0.0000	12.0000	0.0144	0.3415
TOD90	0.0000	12.0000	0.0154	0.3676
TOD91	0.0000	12.0000	0.0120	0.3073
TOD92	0.0000	12.0000	0.0144	0.3403
TOD93	0.0000	12.0000	0.0164	0.3727
TOD94	0.0000	3.0000	0.0004	0.0295
ANFANG1EHE	-2.0000	94.0000	36.1919	34.3342
ANFANG2EHE	-2.0000	94.0000	0.6764	12.8617
ANFANG3EHE	-2.0000	93.0000	-0.5838	9.9655
STATUS1EHE	-2.0000	3.0000	-0.0917	1.5177
STATUS2EHE	-2.0000	3.0000	-1.6783	0.7097
STATUS3EHE	-2.0000	3.0000	-1.7929	0.4149
ENDE1EHE	-2.0000	94.0000	4.9087	21.3192
ENDE2EHE	-2.0000	94.0000	-1.3356	6.0116
ENDE3EHE	-2.0000	85.0000	-1.7720	1.5012

Max number of categories: 16131

Index	SEX	Frequency	Pct	Cumulated	Pct
1	1	7947.00	49.27	7947.00	49.27
2	2	8184.00	50.73	16131.00	100.00

Sum		16131.00	100.00		

Index	GEB	Frequency	Pct	Cumulated	Pct
1	1900	34.00	0.21	34.00	0.21
2	1901	31.00	0.19	65.00	0.40
3	1902	44.00	0.27	109.00	0.68
4	1903	33.00	0.20	142.00	0.88
5	1904	45.00	0.28	187.00	1.16
6	1905	48.00	0.30	235.00	1.46
7	1906	53.00	0.33	288.00	1.79
8	1907	65.00	0.40	353.00	2.19
9	1908	78.00	0.48	431.00	2.67
10	1909	88.00	0.55	519.00	3.22
11	1910	91.00	0.56	610.00	3.78
12	1911	85.00	0.53	695.00	4.31
...					
75	1974	176.00	1.09	15699.00	97.32
76	1975	166.00	1.03	15865.00	98.35
77	1976	147.00	0.91	16012.00	99.26
78	1977	118.00	0.73	16130.00	99.99
79	1979	1.00	0.01	16131.00	100.00

Sum		16131.00	100.00		

Index	KS	Frequency	Pct	Cumulated	Pct

1	1	11603.00	71.93	11603.00	71.93
2	2	4528.00	28.07	16131.00	100.00

Sum		16131.00	100.00		
-----	--	----------	--------	--	--

Index	HEIRAT84	Frequency	Pct	Cumulated	Pct
1	0	15956.00	98.92	15956.00	98.92
2	1	2.00	0.01	15958.00	98.93
3	2	4.00	0.02	15962.00	98.95
4	3	5.00	0.03	15967.00	98.98
5	4	7.00	0.04	15974.00	99.03
6	5	19.00	0.12	15993.00	99.14
7	6	14.00	0.09	16007.00	99.23
8	7	21.00	0.13	16028.00	99.36
9	8	31.00	0.19	16059.00	99.55
10	9	18.00	0.11	16077.00	99.67
11	10	19.00	0.12	16096.00	99.78
12	11	15.00	0.09	16111.00	99.88
13	12	20.00	0.12	16131.00	100.00

Sum		16131.00	100.00		
-----	--	----------	--------	--	--

Index	HEIRAT85	Frequency	Pct	Cumulated	Pct
1	0	15915.00	98.66	15915.00	98.66
2	1	6.00	0.04	15921.00	98.70
3	2	3.00	0.02	15924.00	98.72
4	3	11.00	0.07	15935.00	98.78
5	4	16.00	0.10	15951.00	98.88
6	5	33.00	0.20	15984.00	99.09
7	6	17.00	0.11	16001.00	99.19
8	7	20.00	0.12	16021.00	99.32
9	8	45.00	0.28	16066.00	99.60
10	9	23.00	0.14	16089.00	99.74
11	10	19.00	0.12	16108.00	99.86
12	11	3.00	0.02	16111.00	99.88
13	12	20.00	0.12	16131.00	100.00

Sum		16131.00	100.00		
-----	--	----------	--------	--	--

Index	HEIRAT86	Frequency	Pct	Cumulated	Pct
1	0	15863.00	98.34	15863.00	98.34
2	1	12.00	0.07	15875.00	98.41
3	2	4.00	0.02	15879.00	98.44
4	3	10.00	0.06	15889.00	98.50
5	4	27.00	0.17	15916.00	98.67
6	5	34.00	0.21	15950.00	98.88
7	6	38.00	0.24	15988.00	99.11
8	7	31.00	0.19	16019.00	99.31
9	8	41.00	0.25	16060.00	99.56
10	9	17.00	0.11	16077.00	99.67
11	10	14.00	0.09	16091.00	99.75
12	11	15.00	0.09	16106.00	99.85

13	12	25.00	0.15	16131.00	100.00
----	----	-------	------	----------	--------

Sum		16131.00	100.00		
-----	--	----------	--------	--	--

Index HEIRAT87	Frequency	Pct	Cumulated	Pct
1	0	15917.00	98.67	15917.00 98.67
2	1	8.00	0.05	15925.00 98.72
3	2	18.00	0.11	15943.00 98.83
4	3	5.00	0.03	15948.00 98.87
5	4	13.00	0.08	15961.00 98.95
6	5	20.00	0.12	15981.00 99.07
7	6	25.00	0.15	16006.00 99.23
8	7	19.00	0.12	16025.00 99.34
9	8	28.00	0.17	16053.00 99.52
10	9	19.00	0.12	16072.00 99.63
11	10	21.00	0.13	16093.00 99.76
12	11	29.00	0.18	16122.00 99.94
13	12	9.00	0.06	16131.00 100.00

Sum		16131.00	100.00		
-----	--	----------	--------	--	--

Index HEIRAT88	Frequency	Pct	Cumulated	Pct
1	0	15914.00	98.65	15914.00 98.65
2	1	17.00	0.11	15931.00 98.76
3	2	9.00	0.06	15940.00 98.82
4	3	25.00	0.15	15965.00 98.97
5	4	13.00	0.08	15978.00 99.05
6	5	22.00	0.14	16000.00 99.19
7	6	22.00	0.14	16022.00 99.32
8	7	19.00	0.12	16041.00 99.44
9	8	25.00	0.15	16066.00 99.60
10	9	23.00	0.14	16089.00 99.74
11	10	20.00	0.12	16109.00 99.86
12	11	9.00	0.06	16118.00 99.92
13	12	13.00	0.08	16131.00 100.00

Sum		16131.00	100.00		
-----	--	----------	--------	--	--

Index HEIRAT89	Frequency	Pct	Cumulated	Pct
1	0	15901.00	98.57	15901.00 98.57
2	1	11.00	0.07	15912.00 98.64
3	2	18.00	0.11	15930.00 98.75
4	3	15.00	0.09	15945.00 98.85
5	4	7.00	0.04	15952.00 98.89
6	5	33.00	0.20	15985.00 99.09
7	6	17.00	0.11	16002.00 99.20
8	7	23.00	0.14	16025.00 99.34
9	8	40.00	0.25	16065.00 99.59
10	9	21.00	0.13	16086.00 99.72
11	10	14.00	0.09	16100.00 99.81
12	11	9.00	0.06	16109.00 99.86
13	12	22.00	0.14	16131.00 100.00

Sum		16131.00	100.00		
Index HEIRAT90	Frequency	Pct	Cumulated	Pct	

1	0	15932.00	98.77	15932.00	98.77
2	1	5.00	0.03	15937.00	98.80
3	2	12.00	0.07	15949.00	98.87
4	3	12.00	0.07	15961.00	98.95
5	4	10.00	0.06	15971.00	99.01
6	5	22.00	0.14	15993.00	99.14
7	6	23.00	0.14	16016.00	99.29
8	7	21.00	0.13	16037.00	99.42
9	8	44.00	0.27	16081.00	99.69
10	9	15.00	0.09	16096.00	99.78
11	10	10.00	0.06	16106.00	99.85
12	11	5.00	0.03	16111.00	99.88
13	12	20.00	0.12	16131.00	100.00

Sum		16131.00	100.00		
Index HEIRAT91	Frequency	Pct	Cumulated	Pct	

1	0	15877.00	98.43	15877.00	98.43
2	1	4.00	0.02	15881.00	98.45
3	2	15.00	0.09	15896.00	98.54
4	3	19.00	0.12	15915.00	98.66
5	4	4.00	0.02	15919.00	98.69
6	5	33.00	0.20	15952.00	98.89
7	6	19.00	0.12	15971.00	99.01
8	7	30.00	0.19	16001.00	99.19
9	8	30.00	0.19	16031.00	99.38
10	9	43.00	0.27	16074.00	99.65
11	10	28.00	0.17	16102.00	99.82
12	11	15.00	0.09	16117.00	99.91
13	12	14.00	0.09	16131.00	100.00

Sum		16131.00	100.00		
Index HEIRAT92	Frequency	Pct	Cumulated	Pct	

1	0	15916.00	98.67	15916.00	98.67
2	1	8.00	0.05	15924.00	98.72
3	2	14.00	0.09	15938.00	98.80
4	3	10.00	0.06	15948.00	98.87
5	4	15.00	0.09	15963.00	98.96
6	5	32.00	0.20	15995.00	99.16
7	6	11.00	0.07	16006.00	99.23
8	7	25.00	0.15	16031.00	99.38
9	8	29.00	0.18	16060.00	99.56
10	9	26.00	0.16	16086.00	99.72
11	10	11.00	0.07	16097.00	99.79
12	11	7.00	0.04	16104.00	99.83
13	12	27.00	0.17	16131.00	100.00

Sum 16131.00 100.00

Index	HEIRAT93	Frequency	Pct	Cumulated	Pct
1	0	15894.00	98.53	15894.00	98.53
2	1	12.00	0.07	15906.00	98.61
3	2	8.00	0.05	15914.00	98.65
4	3	15.00	0.09	15929.00	98.75
5	4	26.00	0.16	15955.00	98.91
6	5	27.00	0.17	15982.00	99.08
7	6	19.00	0.12	16001.00	99.19
8	7	22.00	0.14	16023.00	99.33
9	8	29.00	0.18	16052.00	99.51
10	9	35.00	0.22	16087.00	99.73
11	10	13.00	0.08	16100.00	99.81
12	11	9.00	0.06	16109.00	99.86
13	12	22.00	0.14	16131.00	100.00

Sum 16131.00 100.00

Index	HEIRAT94	Frequency	Pct	Cumulated	Pct
1	0	16100.00	99.81	16100.00	99.81
2	1	4.00	0.02	16104.00	99.83
3	2	4.00	0.02	16108.00	99.86
4	3	1.00	0.01	16109.00	99.86
5	4	2.00	0.01	16111.00	99.88
6	5	5.00	0.03	16116.00	99.91
7	6	3.00	0.02	16119.00	99.93
8	7	4.00	0.02	16123.00	99.95
9	8	7.00	0.04	16130.00	99.99
10	9	1.00	0.01	16131.00	100.00

Sum 16131.00 100.00

Index	SCHEID84	Frequency	Pct	Cumulated	Pct
1	0	16095.00	99.78	16095.00	99.78
2	1	4.00	0.02	16099.00	99.80
3	2	3.00	0.02	16102.00	99.82
4	3	2.00	0.01	16104.00	99.83
5	4	3.00	0.02	16107.00	99.85
6	5	3.00	0.02	16110.00	99.87
7	6	1.00	0.01	16111.00	99.88
8	7	3.00	0.02	16114.00	99.89
9	8	3.00	0.02	16117.00	99.91
10	9	5.00	0.03	16122.00	99.94
11	10	4.00	0.02	16126.00	99.97
12	11	3.00	0.02	16129.00	99.99
13	12	2.00	0.01	16131.00	100.00

Sum 16131.00 100.00

Index	SCHEID85	Frequency	Pct	Cumulated	Pct
-------	----------	-----------	-----	-----------	-----

1	0	16101.00	99.81	16101.00	99.81
2	1	3.00	0.02	16104.00	99.83
3	2	4.00	0.02	16108.00	99.86
4	4	2.00	0.01	16110.00	99.87
5	5	5.00	0.03	16115.00	99.90
6	7	2.00	0.01	16117.00	99.91
7	8	3.00	0.02	16120.00	99.93
8	10	3.00	0.02	16123.00	99.95
9	11	3.00	0.02	16126.00	99.97
10	12	5.00	0.03	16131.00	100.00

Sum 16131.00 100.00

Index SCHEID86	Frequency	Pct	Cumulated	Pct
1	0	16090.00	99.75	16090.00
2	1	1.00	0.01	16091.00
3	2	7.00	0.04	16098.00
4	3	2.00	0.01	16100.00
5	4	4.00	0.02	16104.00
6	5	3.00	0.02	16107.00
7	6	3.00	0.02	16110.00
8	7	5.00	0.03	16115.00
9	8	1.00	0.01	16116.00
10	9	2.00	0.01	16118.00
11	10	5.00	0.03	16123.00
12	11	5.00	0.03	16128.00
13	12	3.00	0.02	16131.00

Sum 16131.00 100.00

Index SCHEID87	Frequency	Pct	Cumulated	Pct
1	0	16096.00	99.78	16096.00
2	1	6.00	0.04	16102.00
3	2	5.00	0.03	16107.00
4	3	5.00	0.03	16112.00
5	4	6.00	0.04	16118.00
6	5	3.00	0.02	16121.00
7	6	5.00	0.03	16126.00
8	9	3.00	0.02	16129.00
9	11	2.00	0.01	16131.00

Sum 16131.00 100.00

Index SCHEID88	Frequency	Pct	Cumulated	Pct
1	0	16096.00	99.78	16096.00
2	1	3.00	0.02	16099.00
3	2	1.00	0.01	16100.00
4	4	3.00	0.02	16103.00
5	5	2.00	0.01	16105.00
6	6	4.00	0.02	16109.00
7	7	2.00	0.01	16111.00
8	8	5.00	0.03	16116.00

9	9	5.00	0.03	16121.00	99.94
10	10	3.00	0.02	16124.00	99.96
11	11	5.00	0.03	16129.00	99.99
12	12	2.00	0.01	16131.00	100.00

Sum 16131.00 100.00

Index	SCHEID89	Frequency	Pct	Cumulated	Pct
1	0	16097.00	99.79	16097.00	99.79
2	1	5.00	0.03	16102.00	99.82
3	2	3.00	0.02	16105.00	99.84
4	3	3.00	0.02	16108.00	99.86
5	4	2.00	0.01	16110.00	99.87
6	5	1.00	0.01	16111.00	99.88
7	6	2.00	0.01	16113.00	99.89
8	7	3.00	0.02	16116.00	99.91
9	8	2.00	0.01	16118.00	99.92
10	9	1.00	0.01	16119.00	99.93
11	10	3.00	0.02	16122.00	99.94
12	11	6.00	0.04	16128.00	99.98
13	12	3.00	0.02	16131.00	100.00

Sum 16131.00 100.00

Index	SCHEID90	Frequency	Pct	Cumulated	Pct
1	0	16100.00	99.81	16100.00	99.81
2	1	5.00	0.03	16105.00	99.84
3	3	1.00	0.01	16106.00	99.85
4	4	3.00	0.02	16109.00	99.86
5	5	5.00	0.03	16114.00	99.89
6	6	3.00	0.02	16117.00	99.91
7	7	2.00	0.01	16119.00	99.93
8	9	4.00	0.02	16123.00	99.95
9	10	3.00	0.02	16126.00	99.97
10	11	2.00	0.01	16128.00	99.98
11	12	3.00	0.02	16131.00	100.00

Sum 16131.00 100.00

Index	SCHEID91	Frequency	Pct	Cumulated	Pct
1	0	16086.00	99.72	16086.00	99.72
2	1	3.00	0.02	16089.00	99.74
3	2	8.00	0.05	16097.00	99.79
4	3	4.00	0.02	16101.00	99.81
5	4	4.00	0.02	16105.00	99.84
6	6	4.00	0.02	16109.00	99.86
7	7	1.00	0.01	16110.00	99.87
8	8	3.00	0.02	16113.00	99.89
9	9	2.00	0.01	16115.00	99.90
10	10	6.00	0.04	16121.00	99.94
11	11	5.00	0.03	16126.00	99.97
12	12	5.00	0.03	16131.00	100.00

Sum 16131.00 100.00

Index	SCHEID92	Frequency	Pct	Cumulated	Pct
1	0	16104.00	99.83	16104.00	99.83
2	1	3.00	0.02	16107.00	99.85
3	3	1.00	0.01	16108.00	99.86
4	4	6.00	0.04	16114.00	99.89
5	5	6.00	0.04	16120.00	99.93
6	6	4.00	0.02	16124.00	99.96
7	7	1.00	0.01	16125.00	99.96
8	9	3.00	0.02	16128.00	99.98
9	10	1.00	0.01	16129.00	99.99
10	11	2.00	0.01	16131.00	100.00

Sum 16131.00 100.00

Index	SCHEID93	Frequency	Pct	Cumulated	Pct
1	0	16083.00	99.70	16083.00	99.70
2	1	5.00	0.03	16088.00	99.73
3	2	6.00	0.04	16094.00	99.77
4	3	6.00	0.04	16100.00	99.81
5	4	2.00	0.01	16102.00	99.82
6	5	2.00	0.01	16104.00	99.83
7	6	7.00	0.04	16111.00	99.88
8	7	2.00	0.01	16113.00	99.89
9	8	6.00	0.04	16119.00	99.93
10	9	3.00	0.02	16122.00	99.94
11	10	3.00	0.02	16125.00	99.96
12	11	1.00	0.01	16126.00	99.97
13	12	5.00	0.03	16131.00	100.00

Sum 16131.00 100.00

Index	SCHEID94	Frequency	Pct	Cumulated	Pct
1	0	16114.00	99.89	16114.00	99.89
2	1	7.00	0.04	16121.00	99.94
3	2	3.00	0.02	16124.00	99.96
4	4	3.00	0.02	16127.00	99.98
5	5	2.00	0.01	16129.00	99.99
6	6	1.00	0.01	16130.00	99.99
7	10	1.00	0.01	16131.00	100.00

Sum 16131.00 100.00

Index	TOD84	Frequency	Pct	Cumulated	Pct
1	0	16098.00	99.80	16098.00	99.80
2	2	2.00	0.01	16100.00	99.81
3	3	3.00	0.02	16103.00	99.83
4	4	3.00	0.02	16106.00	99.85
5	5	2.00	0.01	16108.00	99.86

6	6	2.00	0.01	16110.00	99.87
7	7	2.00	0.01	16112.00	99.88
8	8	5.00	0.03	16117.00	99.91
9	9	2.00	0.01	16119.00	99.93
10	10	3.00	0.02	16122.00	99.94
11	11	2.00	0.01	16124.00	99.96
12	12	7.00	0.04	16131.00	100.00

Sum 16131.00 100.00

Index TOD85	Frequency	Pct	Cumulated	Pct	
1	0	16097.00	99.79	16097.00	99.79
2	1	3.00	0.02	16100.00	99.81
3	2	2.00	0.01	16102.00	99.82
4	3	3.00	0.02	16105.00	99.84
5	4	2.00	0.01	16107.00	99.85
6	5	4.00	0.02	16111.00	99.88
7	6	2.00	0.01	16113.00	99.89
8	8	6.00	0.04	16119.00	99.93
9	9	3.00	0.02	16122.00	99.94
10	10	3.00	0.02	16125.00	99.96
11	11	3.00	0.02	16128.00	99.98
12	12	3.00	0.02	16131.00	100.00

Sum 16131.00 100.00

Index TOD86	Frequency	Pct	Cumulated	Pct	
1	0	16085.00	99.71	16085.00	99.71
2	1	5.00	0.03	16090.00	99.75
3	2	2.00	0.01	16092.00	99.76
4	3	4.00	0.02	16096.00	99.78
5	4	3.00	0.02	16099.00	99.80
6	5	3.00	0.02	16102.00	99.82
7	6	5.00	0.03	16107.00	99.85
8	7	3.00	0.02	16110.00	99.87
9	8	7.00	0.04	16117.00	99.91
10	9	4.00	0.02	16121.00	99.94
11	10	4.00	0.02	16125.00	99.96
12	11	2.00	0.01	16127.00	99.98
13	12	4.00	0.02	16131.00	100.00

Sum 16131.00 100.00

Index TOD87	Frequency	Pct	Cumulated	Pct	
1	0	16090.00	99.75	16090.00	99.75
2	1	4.00	0.02	16094.00	99.77
3	2	1.00	0.01	16095.00	99.78
4	3	4.00	0.02	16099.00	99.80
5	4	4.00	0.02	16103.00	99.83
6	5	3.00	0.02	16106.00	99.85
7	6	2.00	0.01	16108.00	99.86
8	7	4.00	0.02	16112.00	99.88

9	8	5.00	0.03	16117.00	99.91
10	9	4.00	0.02	16121.00	99.94
11	10	4.00	0.02	16125.00	99.96
12	11	3.00	0.02	16128.00	99.98
13	12	3.00	0.02	16131.00	100.00

Sum		16131.00	100.00		

Index	TOD88	Frequency	Pct	Cumulated	Pct

1	0	16095.00	99.78	16095.00	99.78
2	1	3.00	0.02	16098.00	99.80
3	3	2.00	0.01	16100.00	99.81
4	4	1.00	0.01	16101.00	99.81
5	5	2.00	0.01	16103.00	99.83
6	6	6.00	0.04	16109.00	99.86
7	7	1.00	0.01	16110.00	99.87
8	8	3.00	0.02	16113.00	99.89
9	9	6.00	0.04	16119.00	99.93
10	10	3.00	0.02	16122.00	99.94
11	11	5.00	0.03	16127.00	99.98
12	12	4.00	0.02	16131.00	100.00

Sum		16131.00	100.00		

Index	TOD89	Frequency	Pct	Cumulated	Pct

1	0	16094.00	99.77	16094.00	99.77
2	1	3.00	0.02	16097.00	99.79
3	2	5.00	0.03	16102.00	99.82
4	3	1.00	0.01	16103.00	99.83
5	4	3.00	0.02	16106.00	99.85
6	5	2.00	0.01	16108.00	99.86
7	6	7.00	0.04	16115.00	99.90
8	7	4.00	0.02	16119.00	99.93
9	8	1.00	0.01	16120.00	99.93
10	9	4.00	0.02	16124.00	99.96
11	10	1.00	0.01	16125.00	99.96
12	11	2.00	0.01	16127.00	99.98
13	12	4.00	0.02	16131.00	100.00

Sum		16131.00	100.00		

Index	TOD90	Frequency	Pct	Cumulated	Pct

1	0	16095.00	99.78	16095.00	99.78
2	1	3.00	0.02	16098.00	99.80
3	2	3.00	0.02	16101.00	99.81
4	3	4.00	0.02	16105.00	99.84
5	4	3.00	0.02	16108.00	99.86
6	5	1.00	0.01	16109.00	99.86
7	7	2.00	0.01	16111.00	99.88
8	8	6.00	0.04	16117.00	99.91
9	9	3.00	0.02	16120.00	99.93
10	10	2.00	0.01	16122.00	99.94

11	11	7.00	0.04	16129.00	99.99
12	12	2.00	0.01	16131.00	100.00

Sum		16131.00	100.00		
-----	--	----------	--------	--	--

Index	TOD91	Frequency	Pct	Cumulated	Pct
1	0	16099.00	99.80	16099.00	99.80
2	1	4.00	0.02	16103.00	99.83
3	2	2.00	0.01	16105.00	99.84
4	3	3.00	0.02	16108.00	99.86
5	4	2.00	0.01	16110.00	99.87
6	5	3.00	0.02	16113.00	99.89
7	6	4.00	0.02	16117.00	99.91
8	7	1.00	0.01	16118.00	99.92
9	8	5.00	0.03	16123.00	99.95
10	9	1.00	0.01	16124.00	99.96
11	10	4.00	0.02	16128.00	99.98
12	11	2.00	0.01	16130.00	99.99
13	12	1.00	0.01	16131.00	100.00

Sum		16131.00	100.00		
-----	--	----------	--------	--	--

Index	TOD92	Frequency	Pct	Cumulated	Pct
1	0	16092.00	99.76	16092.00	99.76
2	1	5.00	0.03	16097.00	99.79
3	2	5.00	0.03	16102.00	99.82
4	3	2.00	0.01	16104.00	99.83
5	4	2.00	0.01	16106.00	99.85
6	5	3.00	0.02	16109.00	99.86
7	6	5.00	0.03	16114.00	99.89
8	7	4.00	0.02	16118.00	99.92
9	8	2.00	0.01	16120.00	99.93
10	9	3.00	0.02	16123.00	99.95
11	10	3.00	0.02	16126.00	99.97
12	11	2.00	0.01	16128.00	99.98
13	12	3.00	0.02	16131.00	100.00

Sum		16131.00	100.00		
-----	--	----------	--------	--	--

Index	TOD93	Frequency	Pct	Cumulated	Pct
1	0	16089.00	99.74	16089.00	99.74
2	1	5.00	0.03	16094.00	99.77
3	2	4.00	0.02	16098.00	99.80
4	3	3.00	0.02	16101.00	99.81
5	4	3.00	0.02	16104.00	99.83
6	5	5.00	0.03	16109.00	99.86
7	6	3.00	0.02	16112.00	99.88
8	7	2.00	0.01	16114.00	99.89
9	8	4.00	0.02	16118.00	99.92
10	9	1.00	0.01	16119.00	99.93
11	10	3.00	0.02	16122.00	99.94
12	11	5.00	0.03	16127.00	99.98

13	12	4.00	0.02	16131.00	100.00

Sum		16131.00	100.00		
Index TOD94	Frequency	Pct	Cumulated	Pct	

1	0	16128.00	99.98	16128.00	99.98
2	1	1.00	0.01	16129.00	99.99
3	2	1.00	0.01	16130.00	99.99
4	3	1.00	0.01	16131.00	100.00

Sum		16131.00	100.00		
Index ANFANG1EHE	Frequency	Pct	Cumulated	Pct	

1	-2	3529.00	21.88	3529.00	21.88
2	-1	3343.00	20.72	6872.00	42.60
3	20	1.00	0.01	6873.00	42.61
4	21	3.00	0.02	6876.00	42.63
5	22	7.00	0.04	6883.00	42.67
6	23	8.00	0.05	6891.00	42.72
7	24	7.00	0.04	6898.00	42.76
8	25	23.00	0.14	6921.00	42.90
9	26	13.00	0.08	6934.00	42.99
...					
64	81	188.00	1.17	15000.00	92.99
65	82	167.00	1.04	15167.00	94.02
66	83	191.00	1.18	15358.00	95.21
67	84	174.00	1.08	15532.00	96.29
68	85	91.00	0.56	15623.00	96.85
69	86	74.00	0.46	15697.00	97.31
70	87	82.00	0.51	15779.00	97.82
71	88	65.00	0.40	15844.00	98.22
72	89	70.00	0.43	15914.00	98.65
73	90	53.00	0.33	15967.00	98.98
74	91	62.00	0.38	16029.00	99.37
75	92	49.00	0.30	16078.00	99.67
76	93	47.00	0.29	16125.00	99.96
77	94	6.00	0.04	16131.00	100.00

Sum		16131.00	100.00		
Index ANFANG2EHE	Frequency	Pct	Cumulated	Pct	

1	-2	11522.00	71.43	11522.00	71.43
2	-1	3258.00	20.20	14780.00	91.62
3	1	619.00	3.84	15399.00	95.46
4	2	160.00	0.99	15559.00	96.45
5	3	62.00	0.38	15621.00	96.84
6	25	1.00	0.01	15622.00	96.84
7	36	1.00	0.01	15623.00	96.85
8	38	2.00	0.01	15625.00	96.86
9	40	1.00	0.01	15626.00	96.87
10	43	1.00	0.01	15627.00	96.88
11	44	3.00	0.02	15630.00	96.89

12	45	4.00	0.02	15634.00	96.92
...					
49	82	17.00	0.11	16055.00	99.53
50	83	35.00	0.22	16090.00	99.75
51	84	23.00	0.14	16113.00	99.89
52	85	4.00	0.02	16117.00	99.91
53	86	1.00	0.01	16118.00	99.92
54	87	1.00	0.01	16119.00	99.93
55	89	2.00	0.01	16121.00	99.94
56	92	1.00	0.01	16122.00	99.94
57	93	8.00	0.05	16130.00	99.99
58	94	1.00	0.01	16131.00	100.00

Sum 16131.00 100.00

Index	ANFANG3EHE	Frequency	Pct	Cumulated	Pct
1	-2	12634.00	78.32	12634.00	78.32
2	-1	3257.00	20.19	15891.00	98.51
3	3	1.00	0.01	15892.00	98.52
4	32	1.00	0.01	15893.00	98.52
5	41	2.00	0.01	15895.00	98.54
6	44	3.00	0.02	15898.00	98.56
7	45	1.00	0.01	15899.00	98.56
8	48	1.00	0.01	15900.00	98.57
9	51	1.00	0.01	15901.00	98.57
10	53	2.00	0.01	15903.00	98.59
11	56	1.00	0.01	15904.00	98.59
...					
34	83	11.00	0.07	16018.00	99.30
35	84	15.00	0.09	16033.00	99.39
36	85	24.00	0.15	16057.00	99.54
37	86	12.00	0.07	16069.00	99.62
38	87	16.00	0.10	16085.00	99.71
39	88	9.00	0.06	16094.00	99.77
40	89	8.00	0.05	16102.00	99.82
41	90	15.00	0.09	16117.00	99.91
42	91	7.00	0.04	16124.00	99.96
43	92	3.00	0.02	16127.00	99.98
44	93	4.00	0.02	16131.00	100.00

Sum 16131.00 100.00

Index	STATUS1EHE	Frequency	Pct	Cumulated	Pct
1	-2	4285.00	26.56	4285.00	26.56
2	-1	3509.00	21.75	7794.00	48.32
3	1	6834.00	42.37	14628.00	90.68
4	2	744.00	4.61	15372.00	95.29
5	3	759.00	4.71	16131.00	100.00

Sum 16131.00 100.00

Index	STATUS2EHE	Frequency	Pct	Cumulated	Pct
-------	------------	-----------	-----	-----------	-----

1	-2	12277.00	76.11	12277.00	76.11
2	-1	3270.00	20.27	15547.00	96.38
3	1	466.00	2.89	16013.00	99.27
4	2	69.00	0.43	16082.00	99.70
5	3	49.00	0.30	16131.00	100.00

Sum		16131.00	100.00		

Index	STATUS3EHE	Frequency	Pct	Cumulated	Pct

1	-2	12829.00	79.53	12829.00	79.53
2	-1	3284.00	20.36	16113.00	99.89
3	1	16.00	0.10	16129.00	99.99
4	2	1.00	0.01	16130.00	99.99
5	3	1.00	0.01	16131.00	100.00

Sum		16131.00	100.00		

Index	EWDE1EHE	Frequency	Pct	Cumulated	Pct

1	-2	11334.00	70.26	11334.00	70.26
2	-1	3311.00	20.53	14645.00	90.79
3	25	1.00	0.01	14646.00	90.79
4	30	1.00	0.01	14647.00	90.80
5	32	1.00	0.01	14648.00	90.81
6	35	2.00	0.01	14650.00	90.82
7	36	1.00	0.01	14651.00	90.83
8	38	2.00	0.01	14653.00	90.84
9	39	4.00	0.02	14657.00	90.86
10	40	7.00	0.04	14664.00	90.91
...					
54	84	82.00	0.51	16084.00	99.71
55	85	20.00	0.12	16104.00	99.83
56	86	3.00	0.02	16107.00	99.85
57	87	1.00	0.01	16108.00	99.86
58	88	2.00	0.01	16110.00	99.87
59	89	1.00	0.01	16111.00	99.88
60	90	4.00	0.02	16115.00	99.90
61	91	5.00	0.03	16120.00	99.93
62	92	7.00	0.04	16127.00	99.98
63	93	3.00	0.02	16130.00	99.99
64	94	1.00	0.01	16131.00	100.00

Sum		16131.00	100.00		

Index	EWDE2EHE	Frequency	Pct	Cumulated	Pct

1	-2	12766.00	79.14	12766.00	79.14
2	-1	3259.00	20.20	16025.00	99.34
3	1	8.00	0.05	16033.00	99.39
4	2	2.00	0.01	16035.00	99.40
5	3	1.00	0.01	16036.00	99.41
6	51	1.00	0.01	16037.00	99.42
7	55	2.00	0.01	16039.00	99.43
8	60	1.00	0.01	16040.00	99.44

9	61	3.00	0.02	16043.00	99.45
10	62	2.00	0.01	16045.00	99.47
11	64	2.00	0.01	16047.00	99.48
...					
29	83	5.00	0.03	16117.00	99.91
30	84	7.00	0.04	16124.00	99.96
31	88	1.00	0.01	16125.00	99.96
32	92	2.00	0.01	16127.00	99.98
33	93	3.00	0.02	16130.00	99.99
34	94	1.00	0.01	16131.00	100.00

Sum		16131.00	100.00		
-----	--	----------	--------	--	--

Index	ENDE3EHE	Frequency	Pct	Cumulated	Pct
1	-2	12859.00	79.72	12859.00	79.72
2	-1	3267.00	20.25	16126.00	99.97
3	70	1.00	0.01	16127.00	99.98
4	81	1.00	0.01	16128.00	99.98
5	82	1.00	0.01	16129.00	99.99
6	83	1.00	0.01	16130.00	99.99
7	85	1.00	0.01	16131.00	100.00

Sum		16131.00	100.00		
-----	--	----------	--------	--	--

End of module tda_ds (63337697).

End of program. Memory used: 67210353 bytes. Tue Sep 10 16:45:17 1996

Anhang D

Statistische Kennzahlen zur Datei *kinder.dat*

In diesem Abschnitt wird zuerst das TDA-commandfile *kinder.cf* gegeben, mit dem die Daten der Datei *kinder.dat* erzeugt worden sind. Als zweiter Teil sind einige statistische Informationen zu dieser Datei aufgelistet.

D.1 *kinder.cf*

```
#-----
section="Einlesen der retrospektiv angegebenen Kinder";
dfile = biolola.dat;
noc = 21747;

PW [8.0](PW9) = c2;
KIND1 <1>[2.0](B0203) = c13;
KIND2 <1>[2.0](B0207) = c16;
KIND3 <1>[2.0](B0211) = c19;
KIND4 <1>[2.0](B0215) = c22;
KIND5 <1>[2.0](B0219) = c25;
KIND6 <1>[2.0](B0223) = c28;
KIND7 <1>[2.0](B0227) = c31;
KIND8 <1>[2.0](B0231) = c34;

#-----
section="Einlesen der Daten aus ppfad.dat";
dfile = ppfad.dat;
noc = 27620;

PW1 [7.0] = c1;
SEX <1>[2.0] = c2;
GEB <2>[4.0] = c3;
KS <1>[1.0] = c4;

match=PW1,PW;

#-----
section"Einlesen der Geburten im 11-Jahres-Fenster ";
dfile = familie.dat;
noc = 27620;

PW2 [7.0] = c1;
GEB84 [0.0] = c71;
GEB85 [0.0] = c72;
GEB86 [0.0] = c73;
GEB87 [0.0] = c74;
GEB88 [0.0] = c75;
GEB89 [0.0] = c76;
GEB90 [0.0] = c77;
GEB91 [0.0] = c78;
GEB92 [0.0] = c79;
GEB93 [0.0] = c80;
GEB94 [0.0] = c81;

match=PW2,PW;

#-----
section"selektieren";
```

```

tsel=eq(KS,1)+eq(KS,2);
keep= PW,SEX,GEB,KS,KIND1,KIND2,KIND3,KIND4,KIND5,KIND6,KIND7,KIND8,
      GEB84,GEB85,GEB86,GEB87,GEB88,GEB89,GEB90,GEB91,GEB92,GEB93,GEB94;
pdata=kinder.dat;

```

D.2 *kinder.dat*

TDA. Analysis of Transition Data (6.0). Fri Feb 16 08:25:35 1996

Command(s): cf=kinder.cf

Section 1. [einlesen retro-Kinder]

Module: tda_gd [data generation]. Current memory: 135593 bytes.

New variables: 9

Set maximum number of cases to: 21747

Will generate a new data matrix with a maximum of 21747 cases.

Data file(s): birolela.dat

Assuming free format data file(s).

Idx	Variable	Label	T	S	WFmt	Definition
1	PW	PW9	3	4	8.0	c2
2	KIND1	B0203	3	1	2.0	c13
3	KIND2	B0207	3	1	2.0	c16
4	KIND3	B0211	3	1	2.0	c19
5	KIND4	B0215	3	1	2.0	c22
6	KIND5	B0219	3	1	2.0	c25
7	KIND6	B0223	3	1	2.0	c28
8	KIND7	B0227	3	1	2.0	c31
9	KIND8	B0231	3	1	2.0	c34

Reading data file: birolela.dat

Read records: 21747

Created data for 9 variable(s) and 21747 cases.

Created a new data matrix.

Records containing (other types of) missing values: 0

End of module tda_gd (396803).

Section 2. [match ppfad vars]

Module: tda_gd [data generation]. Current memory: 396803 bytes.

New variables: 4

Will add variables to the data matrix for 21747 cases.

Data file(s): ppfad.dat

Assuming free format data file(s).

Will read a maximum of 27620 records.

Idx	Variable Label	T	S	WFmt	Definition
1	PN1	3	4	7.0	c1
2	SEX	3	1	2.0	c2
3	GEB	3	2	4.0	c3
4	KS	3	1	1.0	c4

Merging with command: match=PN1,PN
 Sorting variable(s): PN
 Reading data file: ppfad.dat
 Read records: 27620

Created data for 4 variable(s) and 21747 cases.
 Number of cases with no match: 473
 Substituted missing value code: -3

Records containing (other types of) missing values: 0

End of module tda_gd (570854).

Section 3.

Module: tda_gd [data generation]. Current memory: 570854 bytes.

New variables: 12
 Will add variables to the data matrix for 21747 cases.

Data file(s): familie.dat
 Assuming free format data file(s).
 Will read a maximum of 27620 records.

Idx	Variable Label	T	S	WFmt	Definition
1	PN2	3	4	7.0	c1
2	GEB84	3	4	0.0	c71
3	GEB85	3	4	0.0	c72
4	GEB86	3	4	0.0	c73
5	GEB87	3	4	0.0	c74
6	GEB88	3	4	0.0	c75
7	GEB89	3	4	0.0	c76
8	GEB90	3	4	0.0	c77
9	GEB91	3	4	0.0	c78
10	GEB92	3	4	0.0	c79
11	GEB93	3	4	0.0	c80
12	GEB94	3	4	0.0	c81

Merging with command: match=PN2,PN
 Sorting variable(s): PN
 Reading data file: familie.dat
 Read records: 27620

Created data for 12 variable(s) and 21747 cases.
 Number of cases with no match: 473
 Substituted missing value code: -3

Records containing (other types of) missing values: 0

End of module tda_gd (1614971).

Section 4.

Module: tda_cd [changing data structure]. Current memory: 1614971 bytes.

Executing command: tsel=eq(KS,1)+eq(KS,2)

While active, only 16212 cases of the data matrix will be used.

End of module tda_cd (1701977).

Module: tda_pd [writing data/system files]. Current memory: 1701977 bytes.

Command: pdata=kinder.dat

Keep: PW,SEX,GEB,KS,KIND1,KIND2,KIND3,KIND4,
KIND5,KIND6,KIND7,KIND8,GEB84,GEB85,GEB86,GEB87,
GEB88,GEB89,GEB90,GEB91,GEB92,GEB93,GEB94

Data (16212 cases, 23 variables) written to output file: kinder.dat

Separation character: blank

At least one variable is written in free format.

End of module tda_pd (1701977).

Module: tda_ds [descriptive statistics]. Current memory: 1701977 bytes.

Max number of categories: 21747

Index	GEB	Frequency	Pct	Cumulated	Pct
1	1882	1.00	0.01	1.00	0.01
2	1888	1.00	0.01	2.00	0.01
3	1892	6.00	0.04	8.00	0.05
4	1893	2.00	0.01	10.00	0.06
5	1894	3.00	0.02	13.00	0.08
6	1895	5.00	0.03	18.00	0.11
7	1896	13.00	0.08	31.00	0.19
8	1897	11.00	0.07	42.00	0.26
9	1898	10.00	0.06	52.00	0.32
10	1899	29.00	0.18	81.00	0.50
11	1900	34.00	0.21	115.00	0.71
12	1901	31.00	0.19	146.00	0.90
13	1902	44.00	0.27	190.00	1.17
14	1903	33.00	0.20	223.00	1.38
...					
78	1967	392.00	2.42	13871.00	85.56
79	1968	382.00	2.36	14253.00	87.92
80	1969	360.00	2.22	14613.00	90.14
81	1970	305.00	1.88	14918.00	92.02
82	1971	267.00	1.65	15185.00	93.67
83	1972	215.00	1.33	15400.00	94.99
84	1973	204.00	1.26	15604.00	96.25
85	1974	176.00	1.09	15780.00	97.34

86 1975	166.00	1.02	15946.00	98.36
87 1976	147.00	0.91	16093.00	99.27
88 1977	118.00	0.73	16211.00	99.99
89 1979	1.00	0.01	16212.00	100.00

Sum 16212.00 100.00

Index	KS	Frequency	Pct	Cumulated	Pct
1	1	11684.00	72.07	11684.00	72.07
2	2	4528.00	27.93	16212.00	100.00

Sum 16212.00 100.00

Index	SEX	Frequency	Pct	Cumulated	Pct
1	1	7975.00	49.19	7975.00	49.19
2	2	8237.00	50.81	16212.00	100.00

Sum 16212.00 100.00

Index	GEB84	Frequency	Pct	Cumulated	Pct
1	0	15912.00	98.15	15912.00	98.15
2	1	19.00	0.12	15931.00	98.27
3	2	18.00	0.11	15949.00	98.38
4	3	20.00	0.12	15969.00	98.50
5	4	19.00	0.12	15988.00	98.62
6	5	25.00	0.15	16013.00	98.77
7	6	33.00	0.20	16046.00	98.98
8	7	41.00	0.25	16087.00	99.23
9	8	35.00	0.22	16122.00	99.44
10	9	22.00	0.14	16144.00	99.58
11	10	24.00	0.15	16168.00	99.73
12	11	23.00	0.14	16191.00	99.87
13	12	21.00	0.13	16212.00	100.00

Sum 16212.00 100.00

Index	GEB85	Frequency	Pct	Cumulated	Pct
1	0	15879.00	97.95	15879.00	97.95
2	1	34.00	0.21	15913.00	98.16
3	2	32.00	0.20	15945.00	98.35
4	3	17.00	0.10	15962.00	98.46
5	4	21.00	0.13	15983.00	98.59
6	5	35.00	0.22	16018.00	98.80
7	6	40.00	0.25	16058.00	99.05
8	7	14.00	0.09	16072.00	99.14
9	8	28.00	0.17	16100.00	99.31
10	9	38.00	0.23	16138.00	99.54
11	10	24.00	0.15	16162.00	99.69
12	11	34.00	0.21	16196.00	99.90
13	12	16.00	0.10	16212.00	100.00

Sum 16212.00 100.00

Index	GEB86	Frequency	Pct	Cumulated	Pct
1	0	15845.00	97.74	15845.00	97.74
2	1	48.00	0.30	15893.00	98.03
3	2	27.00	0.17	15920.00	98.20
4	3	24.00	0.15	15944.00	98.35
5	4	20.00	0.12	15964.00	98.47
6	5	29.00	0.18	15993.00	98.65
7	6	34.00	0.21	16027.00	98.86
8	7	26.00	0.16	16053.00	99.02
9	8	41.00	0.25	16094.00	99.27
10	9	25.00	0.15	16119.00	99.43
11	10	27.00	0.17	16146.00	99.59
12	11	33.00	0.20	16179.00	99.80
13	12	31.00	0.19	16210.00	99.99
14	85	2.00	0.01	16212.00	100.00

Sum 16212.00 100.00

Index	GEB87	Frequency	Pct	Cumulated	Pct
1	0	15887.00	98.00	15887.00	98.00
2	1	35.00	0.22	15922.00	98.21
3	2	27.00	0.17	15949.00	98.38
4	3	27.00	0.17	15976.00	98.54
5	4	22.00	0.14	15998.00	98.68
6	5	22.00	0.14	16020.00	98.82
7	6	33.00	0.20	16053.00	99.02
8	7	27.00	0.17	16080.00	99.19
9	8	30.00	0.19	16110.00	99.37
10	9	25.00	0.15	16135.00	99.53
11	10	29.00	0.18	16164.00	99.70
12	11	19.00	0.12	16183.00	99.82
13	12	29.00	0.18	16212.00	100.00

Sum 16212.00 100.00

Index	GEB88	Frequency	Pct	Cumulated	Pct
1	0	15835.00	97.67	15835.00	97.67
2	1	34.00	0.21	15869.00	97.88
3	2	35.00	0.22	15904.00	98.10
4	3	32.00	0.20	15936.00	98.30
5	4	35.00	0.22	15971.00	98.51
6	5	40.00	0.25	16011.00	98.76
7	6	19.00	0.12	16030.00	98.88
8	7	25.00	0.15	16055.00	99.03
9	8	31.00	0.19	16086.00	99.22
10	9	23.00	0.14	16109.00	99.36
11	10	31.00	0.19	16140.00	99.56
12	11	29.00	0.18	16169.00	99.73
13	12	43.00	0.27	16212.00	100.00

Sum 16212.00 100.00

Index	GEB89	Frequency	Pct	Cumulated	Pct
1	0	15887.00	98.00	15887.00	98.00
2	1	47.00	0.29	15934.00	98.29
3	2	15.00	0.09	15949.00	98.38
4	3	26.00	0.16	15975.00	98.54
5	4	22.00	0.14	15997.00	98.67
6	5	25.00	0.15	16022.00	98.83
7	6	33.00	0.20	16055.00	99.03
8	7	40.00	0.25	16095.00	99.28
9	8	20.00	0.12	16115.00	99.40
10	9	14.00	0.09	16129.00	99.49
11	10	26.00	0.16	16155.00	99.65
12	11	30.00	0.19	16185.00	99.83
13	12	27.00	0.17	16212.00	100.00

Sum 16212.00 100.00

Index	GEB90	Frequency	Pct	Cumulated	Pct
1	0	15830.00	97.64	15830.00	97.64
2	1	26.00	0.16	15856.00	97.80
3	2	21.00	0.13	15877.00	97.93
4	3	43.00	0.27	15920.00	98.20
5	4	37.00	0.23	15957.00	98.43
6	5	34.00	0.21	15991.00	98.64
7	6	26.00	0.16	16017.00	98.80
8	7	39.00	0.24	16056.00	99.04
9	8	26.00	0.16	16082.00	99.20
10	9	23.00	0.14	16105.00	99.34
11	10	30.00	0.19	16135.00	99.53
12	11	31.00	0.19	16166.00	99.72
13	12	46.00	0.28	16212.00	100.00

Sum 16212.00 100.00

Index	GEB91	Frequency	Pct	Cumulated	Pct
1	0	15892.00	98.03	15892.00	98.03
2	1	29.00	0.18	15921.00	98.21
3	2	27.00	0.17	15948.00	98.37
4	3	26.00	0.16	15974.00	98.53
5	4	21.00	0.13	15995.00	98.66
6	5	27.00	0.17	16022.00	98.83
7	6	26.00	0.16	16048.00	98.99
8	7	41.00	0.25	16089.00	99.24
9	8	23.00	0.14	16112.00	99.38
10	9	28.00	0.17	16140.00	99.56
11	10	23.00	0.14	16163.00	99.70
12	11	28.00	0.17	16191.00	99.87
13	12	21.00	0.13	16212.00	100.00

Sum 16212.00 100.00

Index	GEB92	Frequency	Pct	Cumulated	Pct
1	0	15877.00	97.93	15877.00	97.93
2	1	29.00	0.18	15906.00	98.11
3	2	35.00	0.22	15941.00	98.33
4	3	29.00	0.18	15970.00	98.51
5	4	31.00	0.19	16001.00	98.70
6	5	23.00	0.14	16024.00	98.84
7	6	22.00	0.14	16046.00	98.98
8	7	37.00	0.23	16083.00	99.20
9	8	27.00	0.17	16110.00	99.37
10	9	25.00	0.15	16135.00	99.53
11	10	32.00	0.20	16167.00	99.72
12	11	18.00	0.11	16185.00	99.83
13	12	27.00	0.17	16212.00	100.00

Sum 16212.00 100.00

Index	GEB93	Frequency	Pct	Cumulated	Pct
1	0	15886.00	97.99	15886.00	97.99
2	1	37.00	0.23	15923.00	98.22
3	2	37.00	0.23	15960.00	98.45
4	3	22.00	0.14	15982.00	98.58
5	4	33.00	0.20	16015.00	98.78
6	5	23.00	0.14	16038.00	98.93
7	6	24.00	0.15	16062.00	99.07
8	7	25.00	0.15	16087.00	99.23
9	8	22.00	0.14	16109.00	99.36
10	9	40.00	0.25	16149.00	99.61
11	10	16.00	0.10	16165.00	99.71
12	11	27.00	0.17	16192.00	99.88
13	12	20.00	0.12	16212.00	100.00

Sum 16212.00 100.00

Index	GEB94	Frequency	Pct	Cumulated	Pct
1	0	16138.00	99.54	16138.00	99.54
2	1	26.00	0.16	16164.00	99.70
3	2	25.00	0.15	16189.00	99.86
4	3	12.00	0.07	16201.00	99.93
5	4	3.00	0.02	16204.00	99.95
6	5	1.00	0.01	16205.00	99.96
7	6	2.00	0.01	16207.00	99.97
8	7	3.00	0.02	16210.00	99.99
9	10	2.00	0.01	16212.00	100.00

Sum 16212.00 100.00

Index	KIND1	Frequency	Pct	Cumulated	Pct
1	-2	8590.00	52.99	8590.00	52.99
2	-1	3289.00	20.29	11879.00	73.27

3	13	1.00	0.01	11880.00	73.28
4	19	1.00	0.01	11881.00	73.29
5	20	5.00	0.03	11886.00	73.32
6	21	4.00	0.02	11890.00	73.34
7	22	11.00	0.07	11901.00	73.41
8	23	7.00	0.04	11908.00	73.45
9	24	13.00	0.08	11921.00	73.53
10	25	7.00	0.04	11928.00	73.58
11	26	17.00	0.10	11945.00	73.68

...

69	84	67.00	0.41	16036.00	98.91
70	85	30.00	0.19	16066.00	99.10
71	86	30.00	0.19	16096.00	99.28
72	87	30.00	0.19	16126.00	99.47
73	88	22.00	0.14	16148.00	99.61
74	89	13.00	0.08	16161.00	99.69
75	90	19.00	0.12	16180.00	99.80
76	91	10.00	0.06	16190.00	99.86
77	92	12.00	0.07	16202.00	99.94
78	93	7.00	0.04	16209.00	99.98
79	94	3.00	0.02	16212.00	100.00

Sum 16212.00 100.00

Index	KIND2	Frequency	Pct	Cumulated	Pct
1	-2	9951.00	61.38	9951.00	61.38
2	-1	3284.00	20.26	13235.00	81.64
3	21	1.00	0.01	13236.00	81.64
4	23	2.00	0.01	13238.00	81.66
5	24	4.00	0.02	13242.00	81.68
6	25	6.00	0.04	13248.00	81.72
7	26	10.00	0.06	13258.00	81.78
8	27	5.00	0.03	13263.00	81.81

...

64	83	71.00	0.44	16096.00	99.28
65	84	76.00	0.47	16172.00	99.75
66	85	17.00	0.10	16189.00	99.86
67	86	4.00	0.02	16193.00	99.88
68	87	6.00	0.04	16199.00	99.92
69	88	7.00	0.04	16206.00	99.96
70	89	3.00	0.02	16209.00	99.98
71	90	1.00	0.01	16210.00	99.99
72	92	1.00	0.01	16211.00	99.99
73	93	1.00	0.01	16212.00	100.00

Sum 16212.00 100.00

Index	KIND3	Frequency	Pct	Cumulated	Pct
1	-2	11623.00	71.69	11623.00	71.69
2	-1	3283.00	20.25	14906.00	91.94
3	23	1.00	0.01	14907.00	91.95
4	25	1.00	0.01	14908.00	91.96
5	26	1.00	0.01	14909.00	91.96

6	27	4.00	0.02	14913.00	91.99
7	28	6.00	0.04	14919.00	92.02
8	29	4.00	0.02	14923.00	92.05
...					
61	82	36.00	0.22	16137.00	99.54
62	83	32.00	0.20	16169.00	99.73
63	84	28.00	0.17	16197.00	99.91
64	85	7.00	0.04	16204.00	99.95
65	86	1.00	0.01	16205.00	99.96
66	87	1.00	0.01	16206.00	99.96
67	88	1.00	0.01	16207.00	99.97
68	89	2.00	0.01	16209.00	99.98
69	90	2.00	0.01	16211.00	99.99
70	94	1.00	0.01	16212.00	100.00

Sum		16212.00	100.00		
-----	--	----------	--------	--	--

Index	KIND4	Frequency	Pct	Cumulated	Pct

1	-2	12354.00	76.20	12354.00	76.20
2	-1	3288.00	20.28	15642.00	96.48
3	30	3.00	0.02	15645.00	96.50
4	31	4.00	0.02	15649.00	96.53
5	32	2.00	0.01	15651.00	96.54
6	33	3.00	0.02	15654.00	96.56
7	34	1.00	0.01	15655.00	96.56
8	35	1.00	0.01	15656.00	96.57

...					
52	79	9.00	0.06	16164.00	99.70
53	80	11.00	0.07	16175.00	99.77
54	81	9.00	0.06	16184.00	99.83
55	82	7.00	0.04	16191.00	99.87
56	83	10.00	0.06	16201.00	99.93
57	84	7.00	0.04	16208.00	99.98
58	85	2.00	0.01	16210.00	99.99
59	87	1.00	0.01	16211.00	99.99
60	88	1.00	0.01	16212.00	100.00

Sum		16212.00	100.00		
-----	--	----------	--------	--	--

Index	KIND5	Frequency	Pct	Cumulated	Pct

1	-2	12669.00	78.15	12669.00	78.15
2	-1	3284.00	20.26	15953.00	98.40
3	32	2.00	0.01	15955.00	98.41
4	34	2.00	0.01	15957.00	98.43
5	35	3.00	0.02	15960.00	98.45
6	36	1.00	0.01	15961.00	98.45
7	37	2.00	0.01	15963.00	98.46
8	38	2.00	0.01	15965.00	98.48

...					
44	75	10.00	0.06	16175.00	99.77
45	76	7.00	0.04	16182.00	99.81
46	77	3.00	0.02	16185.00	99.83
47	78	5.00	0.03	16190.00	99.86

48	79	5.00	0.03	16195.00	99.90
49	80	4.00	0.02	16199.00	99.92
50	81	6.00	0.04	16205.00	99.96
51	82	2.00	0.01	16207.00	99.97
52	83	1.00	0.01	16208.00	99.98
53	84	4.00	0.02	16212.00	100.00

Sum 16212.00 100.00

Index	KIND6	Frequency	Pct	Cumulated	Pct

1	-2	12799.00	78.95	12799.00	78.95
2	-1	3281.00	20.24	16080.00	99.19
3	33	1.00	0.01	16081.00	99.19
4	35	1.00	0.01	16082.00	99.20
5	36	2.00	0.01	16084.00	99.21
6	37	2.00	0.01	16086.00	99.22
7	38	1.00	0.01	16087.00	99.23
8	39	2.00	0.01	16089.00	99.24
...					
46	80	5.00	0.03	16207.00	99.97
47	81	4.00	0.02	16211.00	99.99
48	83	1.00	0.01	16212.00	100.00

Sum 16212.00 100.00

Index	KIND7	Frequency	Pct	Cumulated	Pct

1	-2	12864.00	79.35	12864.00	79.35
2	-1	3282.00	20.24	16146.00	99.59
3	38	1.00	0.01	16147.00	99.60
4	39	1.00	0.01	16148.00	99.61
5	40	2.00	0.01	16150.00	99.62
6	41	1.00	0.01	16151.00	99.62
7	42	2.00	0.01	16153.00	99.64
8	44	1.00	0.01	16154.00	99.64
...					
26	70	5.00	0.03	16189.00	99.86
27	71	4.00	0.02	16193.00	99.88
28	72	4.00	0.02	16197.00	99.91
29	73	4.00	0.02	16201.00	99.93
30	74	3.00	0.02	16204.00	99.95
31	75	1.00	0.01	16205.00	99.96
32	76	1.00	0.01	16206.00	99.96
33	81	3.00	0.02	16209.00	99.98
34	82	1.00	0.01	16210.00	99.99
35	83	2.00	0.01	16212.00	100.00

Sum 16212.00 100.00

Index	KIND8	Frequency	Pct	Cumulated	Pct

1	-2	12894.00	79.53	12894.00	79.53
2	-1	3283.00	20.25	16177.00	99.78
3	45	1.00	0.01	16178.00	99.79

4	46	1.00	0.01	16179.00	99.80
5	48	1.00	0.01	16180.00	99.80
6	49	1.00	0.01	16181.00	99.81
7	52	1.00	0.01	16182.00	99.81
8	56	4.00	0.02	16186.00	99.84
9	58	2.00	0.01	16188.00	99.85
10	60	1.00	0.01	16189.00	99.86
11	61	1.00	0.01	16190.00	99.86
12	63	2.00	0.01	16192.00	99.88
13	64	3.00	0.02	16195.00	99.90
14	65	1.00	0.01	16196.00	99.90
15	66	1.00	0.01	16197.00	99.91
16	67	1.00	0.01	16198.00	99.91
17	68	1.00	0.01	16199.00	99.92
18	69	1.00	0.01	16200.00	99.93
19	70	1.00	0.01	16201.00	99.93
20	72	2.00	0.01	16203.00	99.94
21	73	2.00	0.01	16205.00	99.96
22	74	2.00	0.01	16207.00	99.97
23	75	2.00	0.01	16209.00	99.98
24	77	3.00	0.02	16212.00	100.00

Sum		16212.00	100.00		

End of module tda_ds (1701977).

End of program. Memory used: 3964269 bytes. Fri Feb 16 08:27:59 1996

Anhang E

Statistische Kennzahlen zur Datei *ereignisse.dat*

In diesem Abschnitt werden einige statistische Informationen zur Datei *ereignisse.dat*, in der die Ehen- und Kinderereignisse enthalten sind, gegeben. Tabellen, die aus Platzgründen gekürzt werden mußten, sind an den entsprechenden Stellen mit Punkten (...) gekennzeichnet.

TDA. Analysis of Transition Data (6.01). Wed Sep 11 16:28:35 1996

Command(s): cf=ereignisse.cf

Section 1. [Einlesen der Daten]

Module: tda_gd [data generation]. Current memory: 134917 bytes.

New variables: 10

Set maximum number of cases to: 200000

Will generate a new data matrix with a maximum of 200000 cases.

Data file(s): ereignisse.dat

Assuming free format data file(s).

Idx	Variable	T	S	PFmt	Definition
1	PH	3	4	0.0	c1
2	SEX	3	4	0.0	c2
3	GEB	3	4	0.0	c3
4	KS	3	4	0.0	c4
5	JAHR	3	4	0.0	c5
6	MONAT	3	4	0.0	c6
7	EREIGNIS	3	4	0.0	c7
8	ART	3	4	0.0	c8
9	PARTNER	3	4	0.0	c9
10	NR	3	4	0.0	c10

Reading data file: ereignisse.dat

Read records: 26370

Created data for 10 variable(s) and 26370 cases.

Created a new data matrix.

Records containing (other types of) missing values: 0

End of module tda_gd (8135117).

Module: tda_ds [descriptive statistics]. Current memory: 8135117 bytes.

Frequency distribution: freq1=EREIGNIS

Max number of categories: 26370

Index	EREIGNIS	Frequency	Pct	Cumulated	Pct
1	110	11106.00	42.12	11106.00	42.12
2	120	2307.00	8.75	13413.00	50.86
3	199	201.00	0.76	13614.00	51.63
4	300	12756.00	48.37	26370.00	100.00
Sum		26370.00	100.00		

End of module tda_ds (8135117).

=====

Section 2. [Selektion der Eheereignisse]

=====

Module: tda_cd [changing data structure]. Current memory: 8135117 bytes.

Temporary case selection: tsel=EREIGNIS[110,120,199]

While active, only 13614 cases of the data matrix will be used.

End of module tda_cd (8240619).

Module: tda_ds [descriptive statistics]. Current memory: 8240619 bytes.

Frequency distribution: freq1=SEX,GEB,KS,MONAT,JAHR,EREIGNIS,ART,PARTNER,NR
Max number of categories: 26370

Index	SEX	Frequency	Pct	Cumulated	Pct
1	1	6118.00	44.94	6118.00	44.94
2	2	7496.00	55.06	13614.00	100.00

Sum		13614.00	100.00		

Index	GEB	Frequency	Pct	Cumulated	Pct
1	1900	36.00	0.26	36.00	0.26
2	1901	38.00	0.28	74.00	0.54
3	1902	56.00	0.41	130.00	0.95
4	1903	35.00	0.26	165.00	1.21
...					
72	1971	48.00	0.35	13533.00	99.41
73	1972	27.00	0.20	13560.00	99.60
74	1973	31.00	0.23	13591.00	99.83
75	1974	16.00	0.12	13607.00	99.95
76	1975	5.00	0.04	13612.00	99.99
77	1976	2.00	0.01	13614.00	100.00

Sum		13614.00	100.00		

Index	KS	Frequency	Pct	Cumulated	Pct
1	1	10487.00	77.03	10487.00	77.03
2	2	3127.00	22.97	13614.00	100.00

Sum		13614.00	100.00		

Index	MONAT	Frequency	Pct	Cumulated	Pct
1	0	11014.00	80.90	11014.00	80.90
2	1	121.00	0.89	11135.00	81.79
3	2	152.00	1.12	11287.00	82.91
4	3	193.00	1.42	11480.00	84.32
5	4	175.00	1.29	11655.00	85.61
6	5	287.00	2.11	11942.00	87.72

7	6	223.00	1.64	12165.00	89.36
8	7	236.00	1.73	12401.00	91.09
9	8	343.00	2.52	12744.00	93.61
10	9	309.00	2.27	13053.00	95.88
11	10	191.00	1.40	13244.00	97.28
12	11	155.00	1.14	13399.00	98.42
13	12	215.00	1.58	13614.00	100.00

Sum 13614.00 100.00

Index	JAHR	Frequency	Pct	Cumulated	Pct

1	0	210.00	1.54	210.00	1.54
2	1920	1.00	0.01	211.00	1.55
3	1921	3.00	0.02	214.00	1.57
4	1922	7.00	0.05	221.00	1.62
5	1923	8.00	0.06	229.00	1.68
6	1924	7.00	0.05	236.00	1.73
7	1925	22.00	0.16	258.00	1.90
8	1926	13.00	0.10	271.00	1.99
9	1927	14.00	0.10	285.00	2.09

...

67	1985	302.00	2.22	11191.00	82.20
68	1986	343.00	2.52	11534.00	84.72
69	1987	300.00	2.20	11834.00	86.93
70	1988	284.00	2.09	12118.00	89.01
71	1989	303.00	2.23	12421.00	91.24
72	1990	278.00	2.04	12699.00	93.28
73	1991	327.00	2.40	13026.00	95.68
74	1992	267.00	1.96	13293.00	97.64
75	1993	271.00	1.99	13564.00	99.63
76	1994	50.00	0.37	13614.00	100.00

Sum 13614.00 100.00

Index	EREIGNIS	Frequency	Pct	Cumulated	Pct

1	110	11106.00	81.58	11106.00	81.58
2	120	2307.00	16.95	13413.00	98.52
3	199	201.00	1.48	13614.00	100.00

Sum 13614.00 100.00

Index	ART	Frequency	Pct	Cumulated	Pct

1	0	201.00	1.48	201.00	1.48
2	1	9648.00	70.87	9849.00	72.34
3	2	896.00	6.58	10745.00	78.93
4	3	813.00	5.97	11558.00	84.90
5	4	26.00	0.19	11584.00	85.09
6	8	1449.00	10.64	13033.00	95.73
7	9	252.00	1.85	13285.00	97.58
8	10	306.00	2.25	13591.00	99.83
9	21	9.00	0.07	13600.00	99.90
10	22	8.00	0.06	13608.00	99.96

11	23	2.00	0.01	13610.00	99.97
12	24	4.00	0.03	13614.00	100.00

Sum		13614.00	100.00		

Index	PARTNER	Frequency	Pct	Cumulated	Pct

1	0	4805.00	35.29	4805.00	35.29
2	101	1.00	0.01	4806.00	35.30
3	102	1.00	0.01	4807.00	35.31
4	301	1.00	0.01	4808.00	35.32
...					
8240	926802	1.00	0.01	13612.00	99.99
8241	926803	1.00	0.01	13613.00	99.99
8242	927102	1.00	0.01	13614.00	100.00

Sum		13614.00	100.00		

Index	NR	Frequency	Pct	Cumulated	Pct

1	0	201.00	1.48	201.00	1.48
2	1	12465.00	91.56	12666.00	93.04
3	2	903.00	6.63	13569.00	99.67
4	3	43.00	0.32	13612.00	99.99
5	4	2.00	0.01	13614.00	100.00

Sum		13614.00	100.00		

End of module tda_ds (8240619).

=====

Section 3. [Selektion der Eheereignisse]

Module: tda_cd [changing data structure]. Current memory: 8240619 bytes.

Temporary case selection: tsel=EREIGNIS[300]

Current tsel command turned off.

Current number of cases: 26370

While active, only 12756 cases of the data matrix will be used.

End of module tda_cd (8240611).

Module: tda_ds [descriptive statistics]. Current memory: 8240611 bytes.

Frequency distribution: freq1=SEX,GEB,KS,MONAT,JAHR,EREIGNIS,ART,PARTNER,NR
Max number of categories: 26370

Index	SEX	Frequency	Pct	Cumulated	Pct

1	1	1694.00	13.28	1694.00	13.28
2	2	11062.00	86.72	12756.00	100.00

Sum		12756.00	100.00		

Index	GEB	Frequency	Pct	Cumulated	Pct

1	1900	37.00	0.29	37.00	0.29
2	1901	48.00	0.38	85.00	0.67
3	1902	31.00	0.24	116.00	0.91
4	1903	22.00	0.17	138.00	1.08
5	1904	51.00	0.40	189.00	1.48
6	1905	53.00	0.42	242.00	1.90
...					
72	1971	45.00	0.35	12697.00	99.54
73	1972	22.00	0.17	12719.00	99.71
74	1973	19.00	0.15	12738.00	99.86
75	1974	9.00	0.07	12747.00	99.93
76	1975	5.00	0.04	12752.00	99.97
77	1976	3.00	0.02	12755.00	99.99
78	1977	1.00	0.01	12756.00	100.00

Sum		12756.00	100.00		
-----	--	----------	--------	--	--

Index	KS	Frequency	Pct	Cumulated	Pct

1	1	8909.00	69.84	8909.00	69.84
2	2	3847.00	30.16	12756.00	100.00

Sum		12756.00	100.00		
-----	--	----------	--------	--	--

Index	MONAT	Frequency	Pct	Cumulated	Pct

1	0	9292.00	72.84	9292.00	72.84
2	1	365.00	2.86	9657.00	75.71
3	2	299.00	2.34	9956.00	78.05
4	3	278.00	2.18	10234.00	80.23
5	4	264.00	2.07	10498.00	82.30
6	5	284.00	2.23	10782.00	84.52
7	6	292.00	2.29	11074.00	86.81
8	7	318.00	2.49	11392.00	89.31
9	8	283.00	2.22	11675.00	91.53
10	9	264.00	2.07	11939.00	93.60
11	10	264.00	2.07	12203.00	95.66
12	11	272.00	2.13	12475.00	97.80
13	12	281.00	2.20	12756.00	100.00

Sum		12756.00	100.00		
-----	--	----------	--------	--	--

Index	JAHR	Frequency	Pct	Cumulated	Pct

1	1921	1.00	0.01	1.00	0.01
2	1922	8.00	0.06	9.00	0.07
3	1923	5.00	0.04	14.00	0.11
4	1924	12.00	0.09	26.00	0.20
5	1925	14.00	0.11	40.00	0.31
6	1926	22.00	0.17	62.00	0.49
7	1927	19.00	0.15	81.00	0.63

...

67	1987	336.00	2.63	10588.00	83.00
68	1988	391.00	3.07	10979.00	86.07
69	1989	331.00	2.59	11310.00	88.66

70	1990	386.00	3.03	11696.00	91.69
71	1991	323.00	2.53	12019.00	94.22
72	1992	337.00	2.64	12356.00	96.86
73	1993	326.00	2.56	12682.00	99.42
74	1994	74.00	0.58	12756.00	100.00

Sum		12756.00	100.00		

Index	EREIGNIS	Frequency	Pct	Cumulated	Pct

1	300	12756.00	100.00	12756.00	100.00

Sum		12756.00	100.00		

Index	ART	Frequency	Pct	Cumulated	Pct

1	0	12756.00	100.00	12756.00	100.00

Sum		12756.00	100.00		

Index	PARTNER	Frequency	Pct	Cumulated	Pct

1	0	12756.00	100.00	12756.00	100.00

Sum		12756.00	100.00		

Index	NR	Frequency	Pct	Cumulated	Pct

1	1	6170.00	48.37	6170.00	48.37
2	2	3845.00	30.14	10015.00	78.51
3	3	1555.00	12.19	11570.00	90.70
4	4	649.00	5.09	12219.00	95.79
5	5	283.00	2.22	12502.00	98.01
6	6	142.00	1.11	12644.00	99.12
7	7	69.00	0.54	12713.00	99.66
8	8	38.00	0.30	12751.00	99.96
9	9	3.00	0.02	12754.00	99.98
10	10	2.00	0.02	12756.00	100.00

Sum		12756.00	100.00		

End of module tda_ds (8240611).

End of program. Memory used: 9612245 bytes. Wed Sep 11 16:29:07 1996

Anhang F

Literatur

Blossfeld, Hans-Peter, Götz Rohwer (1995): *Techniques of Event History Modelling. New Approaches to Causal Analysis*. Hillsdale, New Jersey. Lawrence Erlbaum Associates.

Landua, Kerstin (1995): *Dokumentation über die Zusammenführung der Biographiedaten des Sozio-ökonomischen Panels*. Berlin.

Rohwer, Götz (1994a): *TDA working papers*. Universität Bremen.

Rohwer, Götz (1994b): *Das Sozio-ökonomische Panel als Datenquelle zur Untersuchung beruflicher Mobilität*. Universität Bremen.